

**A CORPUS LINGUISTIC APPROACH TO MEANING-MAKING
PATTERNS IN SURVEILLANCE DISCOURSE**

by

VIOLA WIEGAND

A thesis submitted to the University of Birmingham for the degree of DOCTOR
OF PHILOSOPHY

Department of English Language and Linguistics
School of English, American and Canadian Studies and Drama
College of Arts and Law
University of Birmingham
August 2019

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

This thesis analyses meaning-making patterns in surveillance discourse using a corpus linguistic approach. As a widespread and contested social issue, surveillance lends itself well to an analysis of meaning in discourse. The thesis puts forward three principles of meaning-making that are explored empirically: (i) meaning evolves with the discourse, (ii) meaning emerges via comparison and (iii) meaning takes shape in co-occurrence patterns. The first principle states that meaning is dynamic and changes across text types and over time. Comparison is therefore necessary to recognise meaning (see principle ii). According to the final principle, meaning can be identified in co-occurrence patterns in discourse. The thesis follows the three principles by taking a comparative approach to co-occurrence patterns of surveillance in corpora that reflect three different social domains: academic discourse, represented by a journal that specialises on surveillance, digital discourse, represented by blog posts that are related to surveillance and, finally, news discourse, represented by a newspaper corpus. The analysis highlights the complexity of surveillance discourses. The thesis develops a methodology that combines traditional corpus linguistic techniques with more qualitative and multimodal elements. By incorporating theoretical frameworks from other disciplines, the thesis demonstrates the interdisciplinary potential of corpus linguistics.

Acknowledgements

I am deeply grateful to my supervisor, Professor Michaela Mahlberg, for her constant encouragement and her guidance that helped me see the light again when I got lost in the data. Likewise, I am thankful for her understanding and flexibility in her other role as my PI and line manager for the CLiC project while I have been finalising this thesis.

I must thank the College of Arts and Law at the University of Birmingham as well as the University of Nottingham for the financial support that made this thesis possible. I would also like to thank Professor Reiner Grundmann and Dr Adam Schembri for their support as part of the supervisory team at different stages of the thesis.

Special thanks to Dr Anthony Hennessey for many insightful discussions, advice on version management and *R*, for the semantic tagging of my corpora, processing the original Times Digital Archive XML, and, especially, for his hard work on the *CorporaCoCo* package. I am grateful for the support given by Dr Simon Branford of the Research Software Group, Advanced Research Computing at the University of Birmingham in running the conversion of the Times Digital Archive data on the High Performance Computing system. I would also like to thank Dr Mike Scott, Dr Andrew Hardie, Professor Laurence Anthony, Dr David Murakami Wood, Sarah Cheung, Dr Andrew Salway, Dr Paul Thompson and Dr Ruth Page, who responded very helpfully to my questions and support requests about tools, data sources and best practices at different stages of this project. I am particularly grateful to Professor Paul Rayson for providing offline versions of the *CLAWS* and *USAS* taggers and answering my questions about the *USAS* tagset.

I am grateful to Dr Anna Čermáková, Professor Wolfgang Teubert and Professor Rodney Jones for insightful discussions and generously sharing their research. A special mention goes out to Dr Magnus Nissel, who suggested the term ‘seed corpus’ (for my analysis in Chapter 5) after my presentation at ICAME 37 in Hong Kong.

I would like to thank my lecturers at the Hong Kong Polytechnic University and the University of Nottingham for inspiring me to attempt this project – and staff at the University of Birmingham for supporting it. I am thankful to my friends Gavin Brookes, Man Yan Leona Keung, Lorenzo Mastropierro, Olumide Popoola and Hang Zhou for proofreading parts of this thesis. Many thanks to Zann Boo for her support and understanding ever since the MA. I am very grateful to my partner Hang Zhou, for his optimism, celebrating every milestone with me, and bearing with me through my self-doubts with this project, while finishing his own thesis and moving on to his post-doc. Many thanks especially for accompanying me for the very final push! And finally, I am grateful to my family, Axel, Beate, and Hendrik Wiegand, who have shown me their unconditional support. Vielen Dank.

Table of contents

1	Introduction.....	1
2	Corpus linguistic approaches to discourse: A linguistic tool kit for studying surveillance.....	10
2.1	Introduction	10
2.2	Surveillance studies and discourse	10
2.3	Meaning in discourse	17
2.4	Corpus linguistic approaches to discourse analysis	21
2.4.1	Theoretical approaches	22
2.4.2	Specialised corpora.....	24
2.4.2.1	Text-external parameters: the situational criterion	24
2.4.2.2	Text-internal parameters: the lexical criterion.....	26
2.4.3	Concordances	27
2.4.4	Collocation and co-occurrence.....	28
2.4.5	Keyness and topicality	32
2.4.6	Semantic tag analysis.....	36
2.5	Discourse domains.....	38
2.5.1	Academic discourse	39
2.5.2	Digital discourse.....	41
2.5.3	News discourse.....	46
2.6	Discourse coordinates	52
2.6.1	Discourse and place	52
2.6.2	Discourse and time	56
2.7	Principles of meaning-making.....	59
3	Methodology: Corpora and analysis frameworks.....	61
3.1	Introduction	61
3.2	Corpora and additional data sources	63
3.2.1	Academic journal articles: The S&S Corpus	66
3.2.1.1	Text-external criteria	67
3.2.1.2	Text retrieval and corpus cleaning.....	70
3.2.1.3	Specifications of the finalised S&S Corpus	73
3.2.2	The Surveillance Blog Corpus	74
3.2.2.1	Text-internal criteria.....	74
3.2.2.2	Text retrieval.....	75

3.2.2.3	Corpus cleaning	79
3.2.2.4	Specifications of the finalised Surveillance Blog Corpus.....	83
3.2.3	The Times Digital Archive, 1986–2008	85
3.2.3.1	The CCR local copy of the complete TDA1785–2008.....	85
3.2.3.2	Processing the data to create the TDA1986–2008 corpus	86
3.2.3.3	Specifications of the finalised TDA1986–2008 corpus	88
3.2.4	Semantically tagged versions of corpora for Chapter 5.....	89
3.2.5	Corpus-external data	90
3.3	Analysis.....	93
3.3.1	Co-occurrence comparisons	93
3.3.2	Implementing co-occurrence comparisons in individual chapters	101
3.3.3	Limitations of co-occurrence comparisons	104
3.4	Conclusions.....	106
4	The academic representation of the concept of surveillance in the <i>Surveillance & Society</i> journal.....	108
4.1	Introduction.....	108
4.2	Defining <i>surveillance</i>	109
4.2.1	Clusters and meaning components of <i>surveillance</i>	111
4.2.2	Explicit definitions of <i>surveillance</i>	118
4.2.2.1	Broad definitions of <i>surveillance</i>	122
4.2.2.2	Specialised definitions of surveillance	126
4.2.2.3	<i>Surveillance</i> and definitions of other concepts	128
4.3	Key keywords in the S&S Corpus	131
4.3.1	Identifying key keywords.....	131
4.3.2	Grouping KKWs.....	133
4.4	Differences between volumes	139
4.4.1	Mehrabov’s (2015) content analysis themes	140
4.4.2	Comparing co-occurrences of KKWs.....	142
4.4.2.1	Volume 1 (2002–2003): The launch.....	145
4.4.2.2	Volume 6 (2009) – The relaunch	150
4.4.2.3	Volume 13 (2015): The final volume of the S&S Corpus	158
4.5	Co-occurrence patterns across the corpus	165
4.5.1	Co-occurrence links and shifts across the volumes	165
4.5.2	A network of surveillance collocates in the S&S Corpus.....	174
4.6	Conclusions.....	181
5	Places and mobility: Interaction orders in the discourse of surveillant landscapes of blog posts.....	183

5.1	Introduction	183
5.2	Parallels between mediated discourse analysis and corpus linguistics.....	186
5.2.1	A mediated discourse analysis of the surveillant landscape in the Bullring shopping centre 189	
5.2.1.1	Discourses in place.....	189
5.2.1.2	Interaction orders	192
5.2.1.3	Historical bodies.....	197
5.2.2	A corpus linguistic analysis of the surveillant landscape in the Bullring shopping centre 200	
5.3	Comparing the co-occurrences of place references in blog posts and journal articles.....	204
5.3.1	Comparing lexical co-occurrences of place nodes	205
5.3.1.1	Co-occurrences of individual place nodes	206
5.3.1.2	Co-occurrences of manually grouped place nodes	211
5.3.2	Co-occurrences of automatically tagged “M” nodes.....	214
5.3.2.1	Co-occurrences of semantic tags for places and air transport	216
5.3.2.2	Co-occurrences of semantically tagged M nodes	224
5.4	Interaction orders of surveillant landscapes in blog posts.....	230
5.4.1	Social aspects	230
5.4.1.1	No obligation or necessity (“S6-”)	234
5.4.1.2	Helping (“S8+”) vs hindering (“S8-”).....	237
5.4.1.3	Power relationships (“S7”)	244
5.4.2	Emotional aspects.....	247
5.4.3	Interactive surveillant landscapes.....	252
5.5	Conclusions.....	253
6	Time: The diachronic representation of surveillance in <i>The Times</i>	256
6.1	Introduction	256
6.2	Long-term co-occurrence patterns.....	259
6.2.1	Long term comparisons for <i>surveillance</i>	261
6.2.1.1	<i>Surveillance</i> collocates in 1986.....	264
6.2.1.2	<i>Surveillance</i> collocates in 1997.....	266
6.2.1.3	<i>Surveillance</i> collocates in 2008.....	269
6.2.2	Long term comparisons for <i>privacy</i>	274
6.2.2.1	<i>Privacy</i> collocates in 1986.....	278
6.2.2.2	<i>Privacy</i> collocates in 1997	279
6.2.2.3	<i>Privacy</i> collocates in 2008.....	283
6.2.3	Long term comparisons for <i>CCTV</i>	284
6.2.3.1	<i>CCTV</i> collocates in 1986	286
6.2.3.2	<i>CCTV</i> collocates in 1997	287
6.2.3.3	<i>CCTV</i> collocates in 2008	288

6.2.4	The development of reflective public discourses on surveillance.....	291
6.3	Co-occurrence patterns at monthly peaks.....	291
6.3.1	Peak comparisons for <i>surveillance</i>	294
6.3.2	Peak comparisons for <i>privacy</i>	301
6.3.3	Peak comparisons for <i>CCTV</i>	305
6.3.4	Peaks supporting and complementing long-term trends	309
6.4	Distribution across sections	310
6.5	Case study: The UK Identity Cards Act 2006.....	314
6.6	Conclusions.....	327
7	Conclusions.....	329
7.1	Introduction.....	329
7.2	Major findings.....	329
7.3	Principles of meaning-making in discourse.....	335
7.4	Limitations	337
7.5	Directions for future work	340
	References	343
	References of cited examples from the S&S Corpus	358
	References of cited examples from the Surveillance Blog Corpus	360
	References of cited examples from the Times Digital Archive.....	363
	Appendix A: List of articles included in the S&S Corpus.....	365
	Appendix B: Articles excluded during the S&S Corpus compilation.....	379
	Appendix C: DEFINE concordance of <i>surveillance</i>	380
	Appendix D: List of issue allocation in Mehrabov's (2015) content analysis*	383
	Appendix E: Volume-based co-occurrence results for Chapter 4.....	384
	Volume 2 vs the whole S&S Corpus.....	384
	Volume 3 vs. the whole S&S Corpus.....	385
	Volume 4 vs. the whole S&S Corpus.....	387
	Volume 5 vs. the whole S&S Corpus.....	388
	Volume 7 vs. the whole S&S Corpus.....	389

Volume 8 vs. the whole S&S Corpus.....	390
Volume 9 vs. the whole S&S Corpus.....	391
Volume 10 vs. the whole S&S Corpus.....	392
Volume 11 vs. the whole S&S Corpus.....	393
Volume 12 vs. the whole S&S Corpus.....	394
Appendix F: List of URLs in the Surveillance Blog Corpus	395

List of Figures

Figure 3-1: Levels of meaning-making in the S&S Corpus and the blog corpus; external vs. lexical criteria for corpus building	64
Figure 3-2: Screenshot illustrating the use of a dash leading to falsely hyphenated words.....	72
Figure 3-3: Screenshot of apparent noise in the ‘t0mear’ blog post (blogspot_outfile369)	82
Figure 3-4: The Times Digital Archive online interface.....	92
Figure 3-5: TDA advanced search for <i>surveillance</i> in 2008, showing the distribution across publication sections	93
Figure 3-6: Flowchart of the co-occurrence comparison method using <i>CorporaCoCo</i>	94
Figure 3-7: Sample output of a TDA corp_text object (extract from monthly_object2008_12.rds).....	96
Figure 3-8: Sample output of a Surveillance Blog Corpus corp_text object (extract from blogspot_outfile056.rds).....	96
Figure 3-9: Illustration of the co-occurrence comparison procedure	98
Figure 3-10: Sample <i>CorporaCoCo</i> plot.....	100
Figure 3-11: Methodological workflow for analysing surveillance discourse.....	106
Figure 4-1: Diagram representing procedure for identifying KKWs in <i>WordSmith Tools</i>	132
Figure 4-2: Co-occurrences of <i>surveillance</i> in Volume 1 vs. the whole S&S Corpus	145
Figure 4-3: Co-occurrences of other KKWs in Volume 1 vs. the whole S&S Corpus.....	146
Figure 4-4: Co-occurrences of <i>surveillance</i> in Volume 6 vs. the whole S&S Corpus	150
Figure 4-5: Co-occurrences of other KKWs in Volume 6 vs. the whole S&S Corpus.....	151
Figure 4-6: Co-occurrences of <i>surveillance</i> in Volume 13 vs. the whole S&S Corpus	160
Figure 4-7: Co-occurrences of other KKWs in Volume 13 vs. the whole S&S Corpus.....	160
Figure 4-8: Significantly different co-occurrences of <i>surveillance</i> in Part I (Volumes 1–5) in comparison to Volumes 6–13	171
Figure 4-9: Significantly different co-occurrences of <i>surveillance</i> in Part II (Volumes 6–13) in comparison to Part I (Volumes 1–5)	172
Figure 4-10: Volume links between the “shared collocates” created with <i>Gephi</i> version 0.9.2	176
Figure 4-11: Shared collocates of <i>data</i>	180
Figure 5-1: Photos from the inside of the Bullring and surveillance signs (taken on 28 July 2018).....	191
Figure 5-2: The beauty stand in Grand Central (photo taken on 28 July 2018).....	194
Figure 5-3: Extract from the Bullring Privacy Policy – “CCTV images”	195
Figure 5-4: The physical signpost promoting the interactive map in the app (photos taken on 28 July 2018)	196
Figure 5-5: Still frames from the Bullring YouTube ad for the “Style Seeker tool” at 48, 49 and 51 seconds.....	196
Figure 5-6: Still frames from the video “SEE EVERYTHING PLUS HAS TO OFFER” on the Bullring & Grand Central website	198
Figure 5-7: Extract from the Bullring Privacy Policy – “Geolocation and Device and Website Use”	199
Figure 5-8: Lexical co-occurrence comparison of <i>space</i> and <i>place</i> in the Surveillance Blog Corpus (left) and the S&S Corpus (right).....	206
Figure 5-9: Co-occurrence comparison of lexical place nodes	209
Figure 5-10: Co-occurrence differences for PLACE nodes (grouping the individual types from 5.3.1.1).....	212

Figure 5-11: Significantly different co-occurrences of the “M7” (“Places”) nodes in the blog posts (left) and journal articles (right) with an effect size over one	219
Figure 5-12: All significantly different co-occurrences of the “M5” (“Movement/transportation: air”) nodes in the blog posts (left) and journal articles (right)	222
Figure 5-13: Co-occurrence comparison of main semantic tags (blog posts left; journal articles right)	226
Figure 5-14: Co-occurrence comparison of the “M” semantic tag and “S” (“Social actions, states and processes”) subcategories	232
Figure 5-15: Co-occurrence comparison of the “M” semantic tag and “E” subcategories....	247
Figure 6-1: The number (upper panel) and percentage (lower panel) of documents in the TDA, 1985–2013, mentioning <i>surveillance</i> ; data retrieved via the TDA “Term Frequency” feature	260
Figure 6-2: Long-term comparisons across the TDA1986–2008.....	261
Figure 6-3: Co-occurrence comparison of <i>surveillance</i> in the year 1986 (left) vs. 1997 (right)	262
Figure 6-4: Co-occurrence comparison of <i>surveillance</i> in the year 1997 (left) vs. 2008 (right)	263
Figure 6-5: Co-occurrence comparison of <i>surveillance</i> in the year 1986 (left) vs. 2008 (right)	264
Figure 6-6: Advert containing <i>surveillance</i> and various 1986 collocates.....	265
Figure 6-7: Advert containing co-occurrences of <i>CCTV</i> with <i>discreet</i> , <i>ends</i> and <i>uncertainty</i>	266
Figure 6-8: Beginning of the response by the Home Secretary to previous letters to the editor	269
Figure 6-9: Raw frequencies of <i>surveillance society</i>	272
Figure 6-10: Co-occurrence comparison of <i>privacy</i> in the whole year 1986 (left) vs. 1997 (right)	275
Figure 6-11: Co-occurrence comparison of <i>privacy</i> in the whole year 1997 (left) vs. 2008 (right)	276
Figure 6-12: Co-occurrence comparison of <i>privacy</i> in the whole year 1986 (left) vs. 2008 (right)	277
Figure 6-13: Example of <i>privacy</i> co-occurring with <i>suit</i> and <i>seeking</i>	278
Figure 6-14: Co-occurrence comparison of <i>CCTV</i> in the whole year 1997 (left) vs. 2008 (right)	285
Figure 6-15: Co-occurrence comparison of <i>CCTV</i> in the whole year 1986 (left) vs. 2008 (right)	285
Figure 6-16: 1986 job ad containing <i>CCTV</i>	286
Figure 6-17: Ad for security services containing <i>CCTV</i>	286
Figure 6-18: Ad for interview skills coaching containing <i>CCTV</i>	287
Figure 6-19: Housing ad containing <i>CCTV</i>	287
Figure 6-20: TV programme description with the co-occurrence pair (<i>CCTV</i> , <i>monitoring</i>).	288
Figure 6-21: Pages 2–3[S] in <i>times2</i> supplementary material of <i>The Times</i> , 21 November 2008	290
Figure 6-22: Relative frequency of <i>surveillance</i> per month, TDA1986–2008	292
Figure 6-23: Relative frequency of <i>privacy</i> per month, TDA1986–2008.....	293
Figure 6-24: Relative frequency of <i>CCTV</i> per month, TDA1986–2008	293
Figure 6-25: Co-occurrence comparison of <i>surveillance</i> in each of the top 5 peak months (vs. 1986), ordered chronologically.....	295

Figure 6-26: Pages 6–7 in <i>The Times</i> , 31 May 2008	300
Figure 6-27: Co-occurrence comparison of <i>privacy</i> in each of the top 5 peak months (vs. 1986), ordered chronologically	302
Figure 6-28: Co-occurrence comparison of <i>CCTV</i> in each of the top 5 peak months (vs. 1999), ordered chronologically	306
Figure 6-29: Beginning of an article from the 2007_03 peak, showing a co-occurrence of <i>CCTV</i> with <i>hotel</i>	308
Figure 6-30: Distribution of documents across newspaper sections in the full TDA1785–2013 (first panel) and containing each of the search terms across the newspaper sections per year	311
Figure 6-31: Frequency profile of <i>id/ identity/ entitlement + cards</i> in the TDA1986–2008.	318

List of Tables

Table 3-1: Overview of RQs	61
Table 3-2: List of S&S issues.....	69
Table 3-3: Distribution of categories in the S&S Corpus	73
Table 3-4: Seed bigrams for the blog corpus	77
Table 3-5: Overview of cleaning procedure for the Surveillance Blog Corpus.....	80
Table 3-6: Breakdown of corpus size	83
Table 3-7: Breakdown of files per blog domain.....	83
Table 3-8: Number of URLs associated with a particular year (excluding the 155 URLs without year indication)	84
Table 3-9: Token counts in the TDA1986–2008.....	88
Table 3-10: Corpus A/B and node words for co-occurrence comparisons Chapters 4–6.....	101
Table 4-1: The 20 most frequent words in the S&S Corpus (# stands for all numbers).....	110
Table 4-2: The 20 most frequent three-word clusters in the <i>surveillance</i> concordance (using the <i>WordSmith Tools</i> Version 7.0.0.167 Index Cluster function “running clusters from a selection” for <i>surveillance</i> , “length 3, maximum working percent 100%, stop at sentence break”).....	111
Table 4-3: KKWs in the S&S Corpus (key in all 13 volumes, compared with the BNC).....	133
Table 4-4: KKWs categorised by <i>USAS</i> tags	134
Table 4-5: KKW groups.....	135
Table 4-6: S&S corpus volumes issues arranged by Mehrabov’s (2015, pp. 119-122) themes	140
Table 4-7: Node sets for the co-occurrence comparison of the S&S Corpus	144
Table 4-8: Shared collocates across the volume comparisons*	167
Table 5-1: Parallels between mediated discourse analysis and a corpus linguistic approach to discourse analysis.....	186
Table 5-2: Top 15 most frequent word types in the Bullring privacy policy generated with <i>AntConc 3.4.4m</i>	200
Table 5-3: Extract from Bennett et al. (2014, p. 87), semantically tagged with the <i>USAS</i> tagger	215
Table 5-4: 20 tokens of the tag “Z2” (“Geographical names”) from the beginning and end of the Surveillance Blog Corpus, respectively.....	217
Table 5-5: The first 20 distinct realisations of each “M” subcategory in Surveillance Blogs Corpus (ordered alphabetically).....	218
Table 5-6: Semantic collocates of “M” nodes with the highest effect sizes (excluding “unique” collocates).....	225
Table 5-7: The first 20 distinct realisations of each “S” subcategory in the Surveillance Blogs Corpus (ordered alphabetically).....	231
Table 5-8: The first 20 distinct realisations of each “E” subcategory in the Surveillance Blogs Corpus (ordered alphabetically).....	247
Table 6-1: Number of documents in the TDA (overall and containing each of the node words)	314
Table 6-2: Important events in the development of the Identity Cards Scheme (adapted from Whitley, 2009/2011, pp. 135–139)	316
Table 6-3: Selection of co-occurrence pairs significantly different in comparison with 2002	322

List of Concordances

Concordance 4-1: Examples of <i>forms of surveillance</i> (17 out of 254) – diversity.....	112
Concordance 4-2: Examples of <i>forms of surveillance</i> (10 out of 254) – controversy.....	112
Concordance 4-3: Examples indicating interdisciplinarity and diversity (20 out of 215 occurrences of <i>of surveillance studies</i>)	114
Concordance 4-4: Examples indicating interdisciplinarity and diversity (21 out of 215 total occurrences of <i>of surveillance studies</i>)	115
Concordance 4-5: Examples of <i>surveillance is a</i> pointing to the diversity of the concept (20 out of 75 total occurrences of <i>surveillance is a</i>).....	116
Concordance 4-6: <i>Surveillance is not</i> + diminishing adjective/adverb + concept (26 out of 79 total occurrences of <i>surveillance is not</i>).....	117
Concordance 4-7: All 18 instances of definitions referring to Lyon’s work	122
Concordance 4-8: Instances problematising broad definitions of surveillance.....	126
Concordance 4-9: Examples of specific definitions of <i>surveillance</i> in the context of (public) health	127
Concordance 4-10: All instances of <i>surveillance</i> as a definition for other issues.....	128
Concordance 4-11: All concordance defining or negotiating the definition of <i>surveillance studies</i>	129
Concordance 4-12: All instances of <i>medical surveillance</i> in Volume 6.....	154
Concordance 4-13: All instances of <i>surveillance medicine</i> in Volume 6.....	155
Concordance 4-14: All instances of (<i>surveillance, revelations</i>) in the S&S Corpus.....	163
Concordance 5-1: All concordance lines of <i>personal data</i> in the Bullring privacy policy....	202
Concordance 5-2: Sample of concordance lines for <i>Bullring</i> from <i>WebCorp</i> with page filter “Birmingham”, within the past six months (accessed on 13 June 2018)	202
Concordance 5-3: All instances of <i>Bullring</i> with page filter <i>security</i> from <i>WebCorp</i> (accessed on 20 June 2018).....	204
Concordance 5-4: <i>Space</i> co-occurring with <i>public, urban</i> or <i>surveillance</i> in the S&S Corpus (five of 732 examples).....	207
Concordance 5-5: All instances of <i>space</i> co-occurring with <i>our</i> in the Surveillance Blog Corpus	207
Concordance 5-6: <i>Place</i> co-occurring with <i>you</i> in the blog posts (ten of 39 examples)	207
Concordance 5-7: All manually grouped PLACE nodes co-occurring with <i>privacy</i> in the blog posts.....	213
Concordance 5-8: Sample of the “M” tag co-occurring with the “L” tag in blog posts (10 out of 3,312 examples)	228
Concordance 5-9: Sample of (“M”, “S7.1-”) co-occurrences in the journal articles (10 out of 648).....	246
Concordance 5-10: Sample of “M” nodes co-occurring with “E3+” (“Calm”) in the blog posts (10 of 388)	249
Concordance 5-11: Sample of “M” nodes co-occurring with “E3-” (“Violent/Angry”) in the blog posts (10 of 2,302).....	249
Concordance 5-12: Sample of “M” nodes co-occurring with “E1” (“Emotion: General”) in the journal articles (10 of 290)	251
Concordance 6-1: All 32 instances of <i>surveillance</i> co-occurring with <i>society</i> in the R1 position in the TDA, 2008	270
Concordance 6-2: All 27 instances of (<i>surveillance, cover</i>)	273
Concordance 6-3: All 7 examples of <i>privacy</i> co-occurring with <i>isolation</i> in 1986.....	279

Concordance 6-4: All 11 examples of <i>privacy</i> co-occurring with <i>law</i> or <i>laws</i> within a span of five and containing <i>princess</i> or <i>princess's</i> in an extended context (span of 15) in the 1997 subcorpus out of 188 total instances	280
Concordance 6-5: All 8 examples of <i>privacy</i> co-occurring with <i>princes</i> in the 1997 subcorpus, shown with an extended context (total span of 15).....	283
Concordance 6-6: All 5 instances of (<i>surveillance</i> , <i>soldiers</i>) in 1988_09	296
Concordance 6-7: All 7 instances of (<i>surveillance</i> , <i>Khan</i>) in 2007_05	297
Concordance 6-8: All 15 instances of <i>privacy</i> co-occurring with <i>civil</i> in 1993_01	303
Concordance 6-9: All 23 instances of (<i>privacy</i> , <i>right</i>) co-occurring in 2008_07	304
Concordance 6-10: All 23 instances of (<i>entitlement</i> , <i>card</i>) in 2002.....	319
Concordance 6-11: Examples of <i>entitlement</i> co-occurring with <i>sense</i> in 2007 out of 32 total instances	321
Concordance 6-12: All 34 instances of <i>id</i> co-occurring with <i>scheme</i> in 2008.....	325

1 Introduction

The theoretical basis for the practical approach to discourse and meaning followed in this thesis views discourse as being at the core of human experience:

The discourse tells us how we can view the world, our private lives, the things we do or don't do, and the things that happen to us. Without the discourse, these things, and even life itself, remain devoid of meaning. (Teubert, 2010, p. 2)

Accordingly, 'discourse' subsumes all acts of communication: all spoken and written texts produced across time. This vast accumulation of language as a whole is not recorded and not available for analysis as a whole. Therefore, we have to make informed decisions on both sampling a specific discourse in a corpus and on the steps for analysing it. The 'meaning' of objects and themes is negotiated in the discourse and evolves with every new contribution.

'Meaning' and 'methods' are two "key notions in corpus linguistics", according to the introduction to the recent volume *The Corpus Linguistics Discourse* (Čermáková & Mahlberg, 2018, p. 3). Both are also central to the present work. Čermáková and Mahlberg (2018) point out that corpus research, including the innovative lexicographical work by Sinclair and colleagues (see e.g. Sinclair, 1987), has given us the insight that meaning is not necessarily found in single words. Sinclair's (2004b) model of the lexical item sees the unit of meaning distributed between the core word and its context. As Krishnamurthy's (2018) contribution to *The Corpus Linguistics Discourse* shows, the question of linking language units to meaning is complex. Nevertheless, in a recent chapter, Teubert (2019) expresses concern at an apparent lack of corpus linguistic studies focusing on meaning. His unease relates mainly to studies that do not focus on discourse, but, for instance, on cognitive linguistic approaches. The body of research on corpus linguistic approaches to discourse has grown extensively in recent years, to the point that one may speak of a subdiscipline that is forming, as apparent from volumes such

as *Corpora and Discourse Studies* (McEnery & Baker, 2015) and *Corpus Approaches to Discourse* (Taylor & Marchi, 2018). These volumes and contributions to other outlets have helped to build up increasingly sophisticated corpus linguistic methodologies for discourse analysis. Perhaps Teubert (2019) still has a point in observing less engagement with the concept of meaning than one might expect. Many corpus linguistic studies of discourse tend to focus on the nature of how a given social issue is represented in the discourse. This work is important, as it often addresses real world concerns of particular social groups. However, these studies do not necessarily focus on conceptualising meaning in discourse. In fact, even the book *Patterns and Meanings in Discourse* (Partington, Duguid, & Taylor, 2013) appears to focus more on case studies of how corpus methods can be used for analysing patterns in discourse rather than conceptualising meaning itself.

The aim of this thesis is to provide a new comparative framework for analysing meaning, which views the meaning of a discourse object as a negotiated and contextual concept. Accordingly, while the foundational definitions of ‘discourse’ and ‘meaning’ closely follow Teubert’s (2010), my approach to analysing meaning in discourse fundamentally relies on comparing texts with different perspectives on the social issue of surveillance. The framework developed in this thesis combines a comparative corpus linguistic methodology with insights from discourse analysis and, to some extent, surveillance studies.

In order to illustrate the workings of meaning in practice, the main research question is: how is surveillance discursively represented? This question is further broken down into specific subquestions (outlined in Section 3.1), because surveillance discourses are complex, as my analysis chapters show. Due to this complexity, surveillance discourses lend themselves well to an analysis of meaning-making patterns. The word *surveillance* itself can refer to many different activities and, depending on the context, is often viewed as contentious. Yet,

surveillance is an important aspect of social life that “has been expanding quietly for many decades and is a basic feature of the modern world” (Bauman & Lyon, 2013, p. vi). As such it is not surprising to see the growing body of research in the interdisciplinary area of surveillance studies. The main concerns of research in this area tend to revolve around issues such as “technical systems, laws, rights and ethics”, as sociolinguist Rodney Jones points out in one of the few linguistic works that engage with surveillance studies (Jones, 2015, p. 408). At the same time, only a few publications in surveillance studies have recognised the potential value of linguistic aspects in a theoretical framework of surveillance (Barnard-Wills, 2011, 2012). Yet, Jones argues that “debates about digital surveillance” have an inherent connection with linguistics, as they “cut to the heart of the most fundamental definitions in our field: what it means to ‘read’, to ‘write’, to ‘speak’, and to ‘listen’” (Jones, 2015, p. 408). In a similar vein, this thesis argues that the onus is not only on surveillance studies scholars to acknowledge the supporting evidence that linguistics can provide. Instead, the analysis of surveillance discourse in relation to the body of work in surveillance studies can in turn inform theories of language and discourse.

This thesis builds on two particular discourse theoretical concepts: first, it operationalises Teubert’s (2010) conceptualisation of discourse and meaning. For the practical implementation to this theoretical approach, the thesis identifies and compares co-occurrence pairs across corpora. Secondly, and more specifically related to surveillance, the thesis develops a corpus linguistic approach to Jones’s (2017) framework of the ‘surveillant landscapes’, which originates from mediated discourse analysis. In addition, the analysis also engages with concepts from the surveillance studies literature, including, in particular, notions of how the modern state of surveillance can be increasingly considered “liquid”, i.e. constantly evolving

(Bauman & Lyon, 2013) and how different actors and actions cooperate to form a ‘surveillant assemblage’ (Barnard-Wills, 2012; based on Haggerty & Ericson, 2000).

The approach to discourse presented here explores both the context and the meaning of the lexical patterns comparatively across different discourse domains. By ‘discourse domains’ I mean subsections of the discourse that differ in their social function, writers and audience. So I view ‘domain’ as a more general concept than ‘genre’ (broadly concurring with Lee, 2002). The purpose of this thesis is to understand the meaning of surveillance as a discourse object, in the usage of different social contexts – academia, digital discourse and news – rather than to distinguish the style of academic writing, blogs and newspaper articles. To begin with, the corpus compiled from the academic journal *Surveillance & Society*, analysed in Chapter 4, contains mostly research articles, but also editorials and book reviews. The blog posts in Chapter 5 reflect a range of textual structures, linguistic features and communicative purposes. Finally, the corpus analysed in Chapter 6 contains the full text of issues from *The Times* from 1986–2008, so mostly news articles, but also, for example, classified advertising and TV programmes.

A project of this scale that aims to both analyse meaning in discourse and to provide a theoretical contribution faces various challenges. Most importantly, it is impossible to retrieve or analyse the entire discourse. So, it is necessary to make a principled selection of texts. However, both meaning and discourse are fuzzy concepts that are not straightforward to translate into analytical steps.

The framework of meaning put forward in this thesis recognises these challenges and addresses them systematically with three ‘principles of meaning-making’:

- (i) meaning evolves with the discourse;
- (ii) meaning emerges via comparison; and
- (iii) meaning takes shape in co-occurrence patterns.

The three principles help to address this complexity by bridging the gap between theory and practice. The principles facilitate an approach to meaning in discourse that is practical enough to facilitate a systematic methodology for an empirical analysis. At the same time, they ensure that the methodology has a theoretical basis. The present study therefore supports the “applied” potential of corpus linguistics and its interdisciplinary contributions.

Principle (i) states that meaning evolves with the discourse, that is, the discourse itself is dynamic. It follows that an empirical analysis has to take informed samples – corpora – from the discourse to have any chance of analysing meaning. However, precisely because of the dynamics of the discourse, it is not enough to take one static sample. Instead, several corpora need to be sampled: identifying which patterns are distinct to a particular domain of the discourse helps to determine in which way they are meaningful. Accordingly, principle (ii) highlights the importance of comparison, which is built into the design of all three studies of this thesis, as co-occurrence patterns are compared across subcorpora and full corpora. In addition, the thesis as a whole is based on a comparative framework, since each analysis chapter focuses on a different discourse domain. Finally, my position in this work is that a useful method of operationalising the identification of meaning in context is to identify co-occurrence patterns (iii). Following these principles, a co-occurrence comparison of real language data is the appropriate method for analysing meaning in discourse.

The concepts that I have touched upon in this introduction – and the common methods of investigating them – are explained in more detail in **Chapter 2**, which reviews the previous

literature in discourse analysis, corpus linguistics and surveillance studies. Chapter 2 also introduces the background to the discourse domains analysed in this thesis: academic discourse, digital discourse and news discourse. The final section discusses the principles of meaning-making in relation to the previous literature.

Chapter 3 elaborates on the methodology of the thesis, which operationalises the principles of meaning-making. The strategy for addressing the problem that discourse is dynamic and context-dependent followed here is to analyse three different perspectives on surveillance discourse through three distinct corpora. The first part of the chapter introduces the three corpora compiled for this study. Two corpora are comparatively specialised in terms of their genre (academic articles and blog posts, respectively) and topic, as both corpora explicitly relate to the concept of surveillance. The third corpus is a large, more general corpus of newspaper articles, extracted from the Times Digital Archive. The second part of Chapter 3 sets out the methodological framework of the analysis, which is based on the novel method of co-occurrence comparisons that colleagues and I have jointly developed. This thesis is the first large-scale study to implement this method and indeed has been a testing ground for various development versions of the *CorporaCoCo R* package released for the method (Hennessey, Wiegand, Mahlberg, Tench, & Lentin, 2017). The basic principle is to count co-occurrences of a node and its collocates in Corpus A and test if this count is significantly different from the count for the same co-occurrence pair in Corpus B. These statistically identified distinct co-occurrences then provide a starting point for a more qualitative analysis of concordances in the corpora.

The core analytical work of the thesis is presented in Chapters 4 to 6, each of which focuses on analysis of surveillance discourse in a specific discourse domain. Starting with the academic domain, **Chapter 4** analyses a specialised corpus of the *Surveillance & Society* (S&S)

journal, the ‘S&S Corpus’. This corpus has been compiled according to ‘external’ criteria, meaning that it includes all research articles, book reviews and editorials until the time of compilation. The analytical emphasis of the chapter is on the meaning of *surveillance* provided by the surveillance scholars. Importantly, this corpus not only represents academic writing but also specialist expert views on surveillance. The chapter analyses the meaning of surveillance from various perspectives. The chapter begins with tracing how scholars characterise and define the meaning of the concept of surveillance. Then, the analysis takes a more lexical approach to the aboutness of this specialised corpus by identifying words that are consistently salient across all volumes of the journal (key keywords). I look for shifts in the way these words are used by analysing their co-occurrence patterns. The chapter shows that the meaning groups formed by the key keywords relate to themes previously identified via a content analysis of the journal (Mehrabov, 2015). However, I argue that the key keywords provide a more high-level representation, whereas the themes are more localised. The chapter concludes with a network of shared collocates of the key keyword *surveillance*, which presents both central topics of the corpus and thematic shifts across the volumes.

Chapter 5 examines meanings of *surveillance* in relation to one of the fundamental coordinates of discourse: place. The chapter draws on the framework of ‘surveillant landscapes’ (Jones, 2017), which originates from the tradition of mediated discourse analysis. The analysis begins with a case study of the multimodal representation of the surveillant landscape in a local shopping centre (chosen as an example of a highly surveilled place). I illustrate how a corpus linguistic approach can complement the qualitative focus of mediated discourse analysis with insights into textual representations of surveillant landscapes more widely. So, the later sections of the chapter apply the surveillant landscape framework to a larger-scale analysis of digital discourse, represented by the Surveillance Blog Corpus. The chapter argues that the social

dimension of surveillant landscapes is particularly important to the framework and also presents the most suitable focus for a corpus linguistic analysis.

Chapter 6 focuses on news discourse in the form of the full data of *The Times* (of London) from 1986–2008 compiled from a local copy of the Times Digital Archive (TDA) for this thesis. Compared to the corpora studied in Chapters 4 and 5, the ‘TDA1986–2008’ corpus is therefore considerably more general – it does not focus on surveillance. The conceptual focus of this chapter is time, which complements the focus on place in Chapter 5 as another fundamental coordinate of discourse. Time is of particular relevance for the negotiation of meaning in the ‘diachronic dimension’ of discourse (Teubert, 2010, p. 210). In practice, the analysis first traces the diachronic development in the frequency and co-occurrence profiles of *surveillance* and the related nodes *privacy* and *CCTV* across the corpus. The chapter shows that the textual location of items within a newspaper issue contributes to their meaning. The final analysis stage explores a specific case study of surveillance discourse that is local to the UK: the representation of the Identity Cards Act 2006, which was debated heatedly and finally repealed. This case study is a valuable example for the linguistic analysis, because politicians, the media and scholars were involved in the debate.

Overall, the focus of the analysis moves from very specialised to more general subsets of discourse. By starting with the experts’ perspective on surveillance, Chapter 4 sets the scene for an analysis of surveillance discourse that would have turned into a distinctly different piece of research if the focus had only been on texts by the general public. Still, the more general perspectives provided by the blog posts and the newspaper in the following chapters are just as important for the analysis and the argument of this thesis. Meaning is found in comparison, not in isolation.

In **Chapter 7**, I bring together the findings from the three analysis chapters. In addition, this chapter brings together the framework of meaning-making principles developed in this thesis as a theoretical and methodological contribution to research on meaning in discourse. The chapter concludes by setting out future directions for research in corpus linguistics and discourse analysis as well as diachronic linguistics and surveillance studies.

2 Corpus linguistic approaches to discourse: A linguistic tool kit for studying surveillance

2.1 Introduction

One of the aims of the present work is to bring together and contrast perspectives on surveillance from various domains of public discourse, including those of the “experts”; surveillance studies scholars. To do this, the thesis draws on the definitions, frameworks, methods and findings from a variety of disciplines, outlined in this chapter. Section 2.2 gives a brief introduction to surveillance studies and links this to recent linguistic research on surveillance discourse. In Section 2.3, I discuss the general theoretical framework for discourse employed in this thesis, followed by a selective survey of corpus linguistic approaches to discourse analysis and the methods relevant to the present work (2.4). The chapter then moves on to introduce the three ‘discourse domains’ analysed in this thesis (2.5): academic discourse, digital discourse and, finally, news discourse. Section 2.6 argues that the external parameters of corpus compilation, place and time, form fundamental coordinates of discourse. Important frameworks and findings from previous research on place and time in discourse are introduced. Finally, Section 2.7 concludes the review of the literature with three principles of meaning-making that form the basis of the present work.

2.2 Surveillance studies and discourse

Surveillance can mean many different things to different people and in different contexts. It is also often considered controversial. As Winseck (2003, p. 187) puts it, “the technological juggernaut driving surveillance must be seen as a double-edged sword”, since the same

technology may be used to disadvantage or to protect people. This tension makes surveillance a suitable theme for studying negotiation in discourse. An obvious example of people being affected by surveillance is privacy invasion, although surveillance also takes other forms. My analysis in Chapter 6 touches upon some arguments on privacy as a social value, for example in relation to the context of the home (see King, 2004, Chapter 2) and paparazzi. In the rest of this section I sketch out some of the fundamental topics in surveillance studies and discuss the emerging linguistic approaches to surveillance.

The surveillance studies literature exhibits a range of conceptualisations of surveillance, as I show in the analysis of a corpus of research articles from the prominent *Surveillance & Society* journal in Chapter 4. Definitions by David Lyon are among the most often cited in the field. In his introductory book *Surveillance Studies: An Overview*, Lyon defines ‘surveillance’ as “the focused, systematic and routine attention to personal details for purposes of influence, management, protection or direction” (Lyon, 2007, p. 14). Yet, the concept is continually evolving. For example, surveillance is moving beyond the domains with which it has been traditionally associated, like crime or terrorism, and is being exploited in order to achieve institutional goals like controlling the health system (see Zurawski, 2007, p. 7).

One well-known type of surveillance technology is CCTV (“closed-circuit television”). The United Kingdom in particular has seen a large growth in the number of CCTV cameras since the 1990s (W. Webster, 2009). They were initially used mostly within retail before the UK government launched open street CCTV campaigns in connection with crime preventions in the 1990s (for a comprehensive overview see Norris, McCahill, & Wood, 2004). In my case study of multimodal surveillant landscapes in Section 5.2, I analyse CCTV announcement signs and other surveillant elements of a local shopping centre. As the following quote points out,

surveillance is a widespread phenomenon in shopping centres (also see Slater, 1998; Walby, 2005).

Video-surveillance has become particularly common in spaces of consumption: shopping malls, the main shopping areas of city centres and inside individual shops. Shopping malls in particular often have an extremely high level of surveillance. (Koskela, 2000, p. 245)

Another particular surveillance scheme that this thesis considers (in Chapter 6) is the introduction of identity cards in the United Kingdom, where they have a contentious history that has been shaped by and reflected in public discourse. The UK government made three attempts in the twentieth and twenty-first centuries to introduce identity cards, all of which eventually failed. Agar (2005) points out that the first two schemes were introduced in war times: first as part of the National Registration Bill in 1915, which aimed at collecting information about British adults in order to determine how many men were available to assist at the front of the First World War. Following the war, the registry was not maintained. However, in the face of the Second World War, Chamberlain's Conservative government introduced the second National Register in 1939 (see Agar, 2005; online). This second scheme was abandoned in the 1950s (by Churchill's Conservative government) after a motorist refused to produce his identity card to police, arguing that they illegitimately made use of wartime policies in peacetime (Agar, 2005). Additional political reasons contributed to the discontinuation of the scheme. Higgs (2004, p. 143) argues that "[t]he identity card scheme was a casualty of the election of a new Conservative government late in 1951". As I explain in Chapter 6, the 2000s saw another attempt to introduce an identity cards scheme, set off by Blair's Labour government. For an overview of how this scheme unfolded, see Whitley (2009/2011). It was eventually repealed by the incoming coalition of the Conservative and Liberal Democrats (e.g. Beynon-Davies, 2011).

Lyon (2009) shows how closely connected the issuing and usage of national identity cards and passports are to social control and state surveillance. Modern ID cards and passports differ from their historical predecessors in two key respects: first, they contain new security features, often including biometrical data such as finger prints and secondly, they are linked to databases which facilitate the collection of further data (see Lyon, 2009, pp. 1–2). Lyon (2009, pp. 8–11) criticises that the terms ‘identity’ and ‘identification’ are often used interchangeably and that ‘ID cards’ are called ‘identity cards’ even though ‘identification’ is the more accurate term. While the two concepts are related (Lyon, 2009, p. 9), ‘identifiers’ as collected by the state with an ID card, for example, can only represent *part* of somebody’s identity (Lyon, 2009, p. 9). Identity is a complex concept and therefore difficult to define or to reduce down to one characteristic (e.g. fingerprint) (also see Hornung & Engemann, 2016).

Importantly, surveillance theory has gradually moved away from traditional models of surveillance that portrayed the source of surveillance as one central actor. One of the influential concepts in this regard is the ‘surveillant assemblage’ (Haggerty & Ericson, 2000), which accounts for a multitude of organisations collecting information and creating ‘data doubles’ of individuals (Haggerty & Ericson, 2000, p. 606). Barnard-Wills (2011, 2012) extends the framework with a particular emphasis on linguistic aspects. He notes the importance of language in framing surveillance, observing that the original concept of an assemblage as introduced by Deleuze and Guattari (1987/2005) involves discourse. In particular, Barnard-Wills (2011, p. 550) argues

[...] that news media discourses form one set of linkages in this surveillant assemblage of enunciation, alongside political language, sales pitches for surveillance technology, signs denoting closed-circuit television (CCTV) coverage, and many other textual forms.

The present work supports this call for more research on the linguistic aspects of surveillance. Like Barnard-Wills (2011), I analyse news media discourses in relation to surveillance – though following a different methodology – in Chapter 6.

In recent years another surveillance theory has gained attention that, like the surveillant assemblage, is based on a wider sociological framework: the concept of ‘liquid surveillance’ (Bauman & Lyon, 2013; Lyon, 2010). It derives from Bauman’s notion of ‘liquid modernity’, which describes the current era of human society, characterised by globalisation and ongoing change:

What was some time ago dubbed (erroneously) ‘postmodernity’, and what I’ve chosen to call, more to the point, ‘liquid modernity’, is the growing conviction that change is the only permanence, and uncertainty the only certainty. (Bauman, 2000/2012, p. viii)

Applied to surveillance, the liquidity metaphor works on various levels. Liquidity can be seen in ‘flows of data’ being transferred across the world in enormous quantities every day, as well as in the ‘time-sensitivity’ of surveillance (Lyon, 2010, p. 325), dependent, for example, on advances in technology and details of security policies being made available for research. In linguistics, the idea of ongoing change seems a rather common-place notion in relation to language. Teubert (2005b, p. 1), for instance, states that “languages are constantly in a state of flux, and the changes they undergo are not predictable”. At the same time, change is an important element of “a pluralist discourse, a discourse in which each member of the discourse society is encouraged to participate in the negotiation of meaning” (Teubert, 2005b, p. 13).

Particularly with regard to recent technological developments in the area of “big data”, it can be said that worldwide surveillance is now not only being transformed in terms of its quantitative extent, but also qualitatively, which raises ethical concerns (Lyon, 2014). Although surveillance scholars were already studying the use of big data for surveillance before Edward

Snowden leaked classified security documents in 2013, the leak raised unprecedented public awareness of these practices (Lyon, 2014).

Lupton (2015, p. 2), explores the ‘digital society’. Apart from activities like digitally reading and participating in (social) media, this includes the “quantified self” movement (Lupton, 2015, p. 181), whereby people are increasingly encouraged to collect and analyse data about their bodies and their lifestyles. Lupton (2015, p. 2) comments:

For some theorists, the very idea of ‘culture’ or ‘society’ cannot now be fully understood without the recognition that computer software and hardware devices not only underpin but actively constitute self-hood, embodiment, social life, social relations and social institutions.

In recent years, some research within surveillance studies has been concerned with discourse (e.g. Barnard-Wills, 2012; Draper, 2012; Kroener, 2013; Schulze, 2015). However, it is fair to say that this has been a minority approach. These studies use a combination of discourse analytical methods, such as “deep reading”, “critical examination” and the analysis of “frames” (Draper, 2012, p. 398) or a “mixed methods analysis, drawing on content and discourse analysis” (Kroener, 2013, p. 121). As applications within an interdisciplinary area, these studies tend to focus on the thematic outcomes of such analyses, rather than on refining linguistic methods and theory.

Several linguistic studies have focused on surveillance as a timely theme following the “Snowden revelations”. Branum and Charteris-Black (2015) investigate part of the British newspaper coverage of the story, using keyword analysis (see Section 2.4.5) to identify differences between reporting by *The Guardian*, *Daily Mail* and *The Sun*. Arguing that the newspapers have different foci in their reporting of the topic, the authors link these characteristics to the notion of news values (see Section 2.5.3). The focus of the study is on the “wider ideological differences between newspapers” (Branum & Charteris-Black, 2015, p. 200), based on the idiosyncratic accounts by the various news outlets, rather than the

characteristics of surveillance discourse in particular. A study of blog posts about surveillance in relation to the Snowden case by Elgesem, Feinerer and Steskal (2016) is mainly motivated by the public attention that the Snowden case drew. The authors argue that the blogs discussing Snowden's story contribute to "civic engagement" (Elgesem et al., 2016, p. 167), although many of the blogs reproduce content from the mainstream news. This suggests that the relations between different public domains of discourse are complex. Looking not at surveillance itself, but the related area of national security, research by MacDonald and colleagues has approached the discourses of national security policies, also using mainly keyword analysis (MacDonald & Hunter, 2013a, 2013b; MacDonald, Hunter, & O'Regan, 2013). Additionally, security is a common topic of interest for linguistics and surveillance studies in the context of language testing policies for immigration. For example, Khan's (2019) ethnographic study of a Yemeni immigrant sets out the linguistic requirements for the UK citizenship process.

Jones has convincingly argued that the concept of surveillance itself has much relevance for linguistics:

[S]cholars of language have a particularly important role to play in discussions of digital surveillance, and, more than that, that understanding practices of surveillance is increasingly central to understanding digital language.

For scholars in other fields, debates about digital surveillance often focus on questions about technical systems, laws, rights and ethics. For linguists, they cut to the heart of the most fundamental definitions in our field: what it actually means to "read," to "write," to "speak," and to "listen." (Jones, 2015, p. 408)

This work has more recently developed into the framework of 'surveillant landscapes', following the principles of mediated discourse analysis, while building upon the body of research in linguistic landscapes. I explain this framework in Section 2.6.1 along with other work on place and discourse.

2.3 Meaning in discourse

Discourse analysis has gained considerable popularity in linguistics (and beyond) in the last two decades, with various approaches being developed. The notion of ‘discourse’ itself is problematic in that it is not only used with different meanings by analysts from numerous disciplines (Biber, Connor, & Upton, 2007) but sometimes also within the work of a particular analyst (Mills, 2004). It is still important to specify which definition of discourse is employed, as the methodology and implications of a study necessarily rely on this.

As an overall approach to the corpus linguistic analysis of discourse, this thesis follows the definition of discourse as conceptualised by Teubert (2010, p. 115), which views discourse at the heart of human experience:

[T]he discourse, and the texts (utterances) that constitute it, are the only firm ground, the only reality we have [...] [T]he reality out there, unmediated by the discourse, including the brute facts of which it consists, is never available to us.

In order to approach discourse empirically, Teubert (2010, p. 116) introduces the distinction between the ‘discourse at large’ and clearly defined subsets, the ‘special discourses’. Accordingly, the discourse at large “consists of all spoken, written or signed utterances from the time when people started using language, in any dialect or language, as long as they had an audience” (Teubert, 2010, p. 116). This definition has two implications: (i) it considers discourse as “language in use”, like the well-known definition by Brown and Yule (1983, p. 1); and (ii) this vast body of language is unavailable for analysis as most of it has not been recorded and is therefore lost. However, it is possible to study the special discourses, as long as there are selection parameters which “define what this selection is in such a way that we can be sure for each text whether it belongs to this particular discourse or not” (Teubert, 2010, pp. 116–117). Even for a narrowly defined special discourse it is often not possible to gather all texts that make up the discourse for inclusion in a corpus. So, the analyst is guided by the research

question when making sampling decisions on which texts to include in the corpus (Teubert, 2010, p. 117). Depending on the research purpose, a discourse may be selected according to “social and institutional contexts” (e.g. newspaper discourse, academic discourse or medical discourse) or “thematic areas” (e.g. the discourse of asylum seekers, surveillance discourse) (Mahlberg, 2014, p. 221).

Teubert (see e.g. 2010, p. 121) argues that there is no ‘discourse-external reality’, meaning that all social concepts are introduced to us via the discourse. This view also implies that no analysis of discourse can identify any objective truth. Instead, meaning is constantly negotiated, it is in flux and therefore prone to change over time. Accordingly, the ‘diachronic dimension’ of discourse is particularly relevant for meaning-making (Teubert, 2010, p. 210). This temporal aspect is the focus of Section 2.6.2 and Chapter 6 in this thesis. As linguists and discourse participants we can trace intertextual links in order to approximate the meaning of lexical items or the ‘discourse objects’ that they represent, i.e. “the concepts of all concrete and abstract things, all properties, all states, actions and processes that are talked about in a discourse” (Teubert, 2010, p. 180).

To look up the meaning of a lexical item we often turn to a dictionary. Corpus linguistics has had an important impact on dictionary making, including aspects such as which words should be included, what additional information should be added (e.g. examples of real usage) and, importantly, how definitions should be compiled. While corpus methods have therefore helped to improve the relevance and usability of dictionaries, dictionary definitions, by their very nature, are still limited when it comes to capturing meaning:

A dictionary definition should ideally be a generalization of the meaning of a given lexical item. It should describe the lexical item as a type, as a common denominator of all the different occurrences, which it represents. It should be compatible with all (or most of the) occurrences of the item in the discourse. But such a definition does not capture what a particular occurrence of this lexical item means. (Teubert, 2007b, p. 68)

Dictionary definitions are therefore produced in a similar process to the corpus linguistic approaches to discourse analysis that I describe in Section 2.4. However, as Teubert's (2007b) quote shows, dictionaries have a more general purpose than discourse studies which often focus on a specific context rather than generalising over all occurrences of a particular word across the language.

There has been disagreement on the role of an individual's mind in meaning-creation, in addition to meanings created in the discourse itself. Teubert (2010, p. 42) maintains that we "do not have direct access to the mind". For him, meaning can only be found in the discourse, where we can interpret it with linguistic techniques. However, on the opposite end of the spectrum, Langacker (2008, p. 27) argues:

From a cognitive linguistic perspective, the answer is evident: meanings are in the minds of the speakers who produce and understand the expressions. It is hard to imagine where else they might be.

Many linguists appear to take a view that is located between these two extremes. Sinclair acknowledges both the role that individual minds play in the creation of meaning and the problem of accessing these minds. Interviewed by Teubert as part of the 2004 publication of the OSTI report¹, Sinclair suggests that "meaning is an impression in the mind of an individual, and that is impenetrable, using linguistic techniques" (Sinclair, Jones, & Daley, 1970/2004, p. xxviii). Similarly, Mahlberg (2005, p. 188) argues that "[m]eaning has a subjective dimension that refers to individual language experience, and meaning has a social dimension that manifests its role in the social reality of the discourse community". One corpus linguistic framework that focuses on the textual effects on individual minds is Hoey's (2005) theory of 'lexical priming', which suggests that language users are primed towards certain contextual uses of particular

¹ "OSTI" stands for the UK Government Office for Scientific and Technical Information, for which the report was originally produced.

words. In stylistics, where the role of the reader of a text is of particular importance, a body of research has developed that combines cognitive frameworks and methods with the analysis of textual evidence. For instance, Culpeper's (2001, pp. 35–36) model of characterisation combines “top-down processes (that is, determined by knowledge in memory) and bottom-up processes (determined by textual elements)”. A growing body of work in “reader response research” (for an overview see Whiteley & Canning, 2017) seeks to empirically study how readers react to features of texts.

Among the text-focused approaches to meaning, one of the methods of deriving meaning is comparison. In stylistics, this approach is conceptualised in terms of identifying ‘deviations’ from linguistic norms (Leech, 1985; Mahlberg & Wiegand, 2018). In corpus linguistics, comparison is similarly important for identifying meaning (Mahlberg, 2007b; McEnery & Xiao, 2010; on diachronic meaning change, see Koteyko, 2007). Indeed, all corpus linguistic methods can be considered “inherently comparative” (McEnery & Xiao, 2010, p. 175; also see Mahlberg, 2013, p. 24; Tognini-Bonelli, 2001, p. 139). Mahlberg (2014, p. 223) points out that comparative approaches are often particularly important for the study of specific discourses.

Another tradition of discourse analysis that this thesis draws on selectively is known as ‘mediated discourse analysis’. This approach to discourse analysis differs from the way that discourse is generally conceptualised in corpus linguistic studies in that it emphasises the role of multiple semiotic modes in meaning-making. Rather than a specific text, it takes a ‘mediated action’ as the unit of analysis (Jones, 2012, p. 28), e.g. the action of buying a cup of coffee, including all involved participants and material objects. Accordingly, mediated discourse analysis is not only concerned with written and spoken text, but also with other semiotic sources. Within this tradition, a framework for analysing social actions has been introduced by

Scollon and Scollon (2004; also see Scollon, 2001; Scollon & Scollon, 2003). This framework (also called ‘nexus analysis’) broadly views social actions as the ‘nexus’ of three elements:

A social action takes place as an intersection or nexus of some aggregate of discourses (educational talk, for example) – the *discourses in place*, some social arrangement by which people come together in social groups (a meeting, a conversation, a chance contact, a queue) – *the interaction order*, and the life experiences of the individual social actors – *the historical body*. (Scollon & Scollon, 2004, p. 19; emphasis in original)

The notion of ‘surveillant landscapes’ that forms the focus of the analysis in Chapter 5 builds on these three elements of social actions. It specifically focuses on how people interact with the surveillant elements in their environment as they go about their daily life (see Section 2.6.1; Jones, 2017).

2.4 Corpus linguistic approaches to discourse analysis

Corpus linguistic methods have received much attention in recent decades. In some linguistic sub-disciplines, commentators have talked of a “corpus turn” (e.g. Kleinke, 2012, p. 424 in pragmatics; Leech & Short, 2007, p. 286 in stylistics). Already in the early nineties, M. Baker (1993, p. 233) argued that “[t]he rise of corpus linguistics has serious implications for any discipline in which language plays a major role”. Essentially, analyses in the diverse field of corpus linguistics examine texts with the help of software tools in order to answer particular research questions (McEnery & Hardie, 2012). As corpus linguistic methods allow researchers to empirically analyse large quantities of language data both quantitatively and qualitatively (P. Baker, 2010), they have been applied in a wide range of contexts.

Corpus linguistic approaches to discourse analysis have gained considerable popularity in recent years. It has been suggested that researchers in the tradition of discourse analysis per se and “pure” corpus linguistics “set out to describe and explain very different realities, sustain very different views of what constitutes evidence, and have different views of the kinds of

claims that can be made” (Virtanen, 2009, p. 62). Yet, the two areas share an interest in genuine language use (Mahlberg, 2014; Virtanen, 2009) and there is an increasing number of studies that successfully examine “the representation of social issues, global events or groups in society” with corpus linguistic tools (Mahlberg, 2014, p. 220).

In the following subsections, I first present various theoretical views on using corpus approaches to discourse analysis (2.4.1) and then introduce research on specialised corpora (2.4.2). The later subsections each cover a particular methodological aspect relevant to this thesis: concordances (2.4.3), collocation (2.4.4), keyness (2.4.5) and, finally, semantic tagging (2.4.6).

2.4.1 Theoretical approaches

The theoretical view taken in corpus linguistic approaches naturally depends on the researcher’s understanding of meaning and discourse. As shown in Section 2.3, a variety of views exist. Work in the Sinclairian tradition tends to be associated with an approach that assigns a theoretical status to the “field” of corpus linguistics. However, it is also possible to use corpus linguistic methods to test theories developed in other fields (Hardie & McEnery, 2010), not only in terms of cognitive approaches (see Section 2.3), but for example theoretical frameworks from fields like history or sociology.

Sinclair (e.g. 2004b, p. 141) focuses on inter-word relationships, conceptualised in his model of ‘co-selection’ which involves various types of co-occurrence patterns around a central ‘lexical item’ (which could be a word or a phrase). Cheng (2006; also see Cheng & Lam, 2010) has applied this model in order to investigate media discourses related to Hong Kong. In a similar spirit to Sinclair’s contributions, several other corpus linguists have developed theoretical approaches based on corpus methods and findings. These include Teubert’s (2010)

work on meaning negotiation based on corpus evidence and Hoey's (2005) theory of lexical priming (see Section 2.3). Mahlberg (2014, p. 217, also see 2005, pp. 188–189) presents three tenets of the 'corpus theoretical approach': (i) "[l]anguage is a social phenomenon"; (ii) "[m]eaning and form are associated"; and (iii) "[a] corpus linguistic description of language prioritises lexis".

An increasing body of corpus linguistic studies is associated with critical discourse analysis. Corpus linguistic methods have been employed to work with theoretical frameworks from other traditions within and beyond linguistics in order to analyse discourse. For example, McEnery (2006) effectively builds on the sociological framework of moral panic theory to analyse writings about immorality and bad language in the twentieth century. In some studies, corpus linguistic methods contribute to the research questions of another discipline by shedding light on the language use for a particular dataset. An example of such a distinctly "applied" use of corpus methods is the analysis of questionnaire data (e.g. in the fields of psychology of religion; Altmeyer et al., 2015; veterinary science; Huntley et al., 2018; and anthropology; Nolte, Ancarno, & Jones, 2018).

Although many corpus studies of discourse apply the same basic tools that are introduced in the following sections, Marchi and Taylor (2018, p. 5) point out that

rather than one approach there is, in fact, a diverse range of kin approaches that go under different names such as *corpus-based CDA*, (e.g. Baker et al. 2008), *CADS* (Partington 2004), *discourse-oriented corpus studies* (Gabrielatos, private conversation), *corpora and discourse studies* (Baker & McEnery 2015) or under no particular label. [emphasis in original]

The view taken in this thesis agrees with Marchi and Taylor's (2018, p. 5) aim "to overstep disciplinary barriers and avoid pigeon-holing or branding" and therefore does not seek to prefer one of these labels over the others. Instead, the thesis is grounded in the definition of discourse

as explained in Section 2.3 and also draws on additional frameworks – such as the notion of surveillant landscapes which are introduced in 2.6.1 – as appropriate for the research objectives.

2.4.2 Specialised corpora

The value of the findings of any study depends on the suitability of the data. Studies concerned with a particular discourse type often require the compilation of specialised corpora (Partington et al., 2013). The parameters for corpus compilation are tied to the research question(s) and “define” the discourse that a specific corpus is meant to represent (Teubert, 2010, p. 116). Transparency is crucial in reporting these parameters (Sinclair, 2005; Teubert, 2010) in order to allow others to interpret the contribution of the study and to replicate the results if possible. The “parameters can be either text-internal (e.g. language, or the occurrence of certain lexical items), or text-external, like parameters of space or time or situation, or categories describing people” (Teubert, 2010, p. 117). The following two subsections provide examples of corpora with these compilation parameters, although some use a combination of both.

2.4.2.1 Text-external parameters: the situational criterion

One of the basic text-external parameters of a corpus is authorship. Unlike general corpora, a specialised corpus may contain only writing by one particular author or a restricted group of authors. This is often the case in corpus stylistic research. For example, Mahlberg and McIntyre (2011) analyse the representation of the fictional world in Fleming’s novel *Casino Royale*. Mastropierro (2017) studies differences in the stylistic effects between Joseph Conrad’s *Heart of Darkness* and four of its Italian translations.

Corpora used for discourse analysis are rarely based on texts from a single author, except in special cases where the focus is on texts from a prominent figure such as speeches by a

particular politician (see e.g. Charteris-Black, 2014), although even then multiple people may have been involved in drafting the texts. More commonly, the concept of authorship in specialised corpora compiled for discourse analysis is not tied to an individual but to an institution. For instance, Teubert (2007a) analyses the discourse of the catholic social doctrine. Focusing on corporate communication, Turnbull (2013, p. 294) analyses the “frequently asked questions” website of a large consumer product manufacturer.

Time is a common parameter for corpus compilation, in particular in the case of media discourse. For instance, Mahlberg (2007c) analyses a corpus of all articles from *The Guardian* published in the year 2002 that contain the phrase *sustainable development*. By dividing the corpus into monthly subcorpora, Mahlberg was able to study the usage of that phrase in relation to events across that year. An example of a corpus built on the parameter of place is Busse’s (2019) corpus of interviews with residents of the borough of Brooklyn, New York City.

For the compilation of specialised corpora, it has been suggested to consult “experts in the field” for advice on what to include in order to make these discourses as representative as possible (Cheng, 2012, p. 166). The parameters that researchers consider to compile corpora to some extent overlap with the ‘situational characteristics’ in register analysis (e.g. Biber & Conrad, 2009, Chapter 2). In a similar way to Cheng (2012), Biber and Conrad (2009, p. 38) recommend that corpus compilers obtain information from “expert informants” in order to understand what factors are important for a particular register. Register analysis considers the physical and social circumstances in which a text is produced, in a similar manner as mediated discourse analysis focuses on actions (see Section 2.3). However, register analysis takes the text as a unit of meaning and does not include material objects in the analysis. In Chapter 5, I argue that this textual unit of meaning is the basis of corpus linguistics in general, but that this approach is complementary to mediated discourse analysis.

2.4.2.2 Text-internal parameters: the lexical criterion

Words and phrases – and the patterns that they form – are the basis of corpus linguistic methods. This foundation is spelled out in Mahlberg’s (2005, p. 189) corpus theoretical approach (see Section 2.4.1): “[a] corpus linguistic description of language prioritises lexis”. That is to say that lexical co-occurrence patterns – rather than the structures described in traditional grammars – form the basis of meaning of any text. Accordingly, text-internal selection parameters for specialised corpora often relate to lexis, for example in the form of particular search terms. The implied assumption tends to be that if candidate texts contain certain words, then they are thematically associated with the concepts represented by the words.

Depending on the discourse domain, different procedures are followed. The prototypical example for this approach is the retrieval of newspaper articles from a database, which is a popular source of data for corpus linguistic studies of discourse (see e.g. Gabrielatos & Baker, 2008; McEnery, McGlashan, & Love, 2015; Schröter & Storjohann, 2015; also see Section 2.5.3). For genres that are not readily available from an individual database, other approaches have to be taken. Research in the web-as-corpus tradition, for instance, sometimes employs a so-called ‘bootstrapping’ procedure to collect URLs from the web that contain a set of ‘seed’ words. I take this approach for one of the corpora in this thesis (see Section 3.2.2), using the *BootCaT* tool (Baroni & Bernardini, 2004). As with the search terms for the newspaper articles, these words are chosen based on their thematic suitability. Depending on the required size of the corpus, researchers may extract salient words from the collected texts as additional seeds to run another cycle of text collection (for an explanation of the procedure also see Gatto, 2014).

While the use of search terms is widespread for collecting newspaper articles and other texts from databases and websites, the issue of search term selection has been approached mainly from an operational perspective. For example, Gabrielatos (2007) puts forward the

‘relative query term relevance’ measure that allows the researcher to estimate how the relevance of a query can be improved by adding more search terms to the set of ‘core query terms’. Gabrielatos (2007) argues that while a concept can be referred to indirectly without the text containing the exact core query, selecting enough relevant – and only relevant – search terms is a balancing act.

An alternative approach to choosing search terms might be to focus on intertextual links between documents. Teubert (2019) makes a case for a future corpus tool that allows discourse participants beyond academia to get access to the linguistic methods for analysing meaning. In this scenario, anybody can engage with evidence from a large, constantly updating corpus in an interface that displays intertextual links and paraphrases, effectively showing meaning negotiation of particular lexical items in progress. While this vision appears to move away from the discourse analytic focus on a specialised corpus, it has as yet not been put into practice.

2.4.3 Concordances

A basic corpus linguistic tool is the concordance, which shows a vertical listing of all occurrences of a word or phrase surrounded by the immediate context (e.g. McEnery & Hardie, 2012). In his seminal book *Corpus, Concordance, Collocation*, Sinclair (1991, p. 42) argues that “[t]he quality of evidence about the language which can be provided by concordances is quite superior to any other method”. The analysis of concordances makes it possible to recognise patterns that would not otherwise be noticeable, for example through the help of a “sort” function in a concordancer like *WordSmith Tools* (Scott & Tribble, 2006). As Sinclair (1991, p. 100) points out, “[t]he language looks rather different when you look at a lot of it at once”. Accordingly, the concordance is an important type of visualisation in corpus linguistics

(see Mahlberg & Wiegand, 2018, sec. 3). At its most basic, the concordance helps to provide evidence for the “existence” of a pattern (Wynne, 2008, p. 708).

Increasingly, concordance lines are not only used as a first step in a concordance analysis, but also as a follow-up procedure for analysing the context of words and phrases that have been highlighted by statistical methods. While Sinclair (1991) emphasises the need for studying a complete set of concordance lines, when possible, or to follow systematic selection criteria when necessary, Jeffries and Walker (2018, p. 197) criticise the fact that corpus studies tend to be vague about these criteria.

2.4.4 Collocation and co-occurrence

Collocation is a lexical pattern and a central concept in corpus linguistics, but has been defined in many different ways (McEnery & Hardie, 2012, Chapter 6). Among the most prominent figures in collocation research are Firth and Sinclair. Approaches building on their work have been referred to as “Firthian” (Evert, 2008, p. 1220) or “neo-Firthian” (McEnery & Hardie, 2012, p. 122). Collocation is one of the five ‘categories of co-selection’ that form Sinclair’s (2004b) ‘lexical item’, a framework which emphasises that meaning is created beyond the single word in a text. Sinclair (2004b, p. 141) defines collocation as “the co-occurrence of words with no more than four intervening words”. Usually, researchers are interested in recurrent co-occurrences. In combination with wider lexicogrammatical patterns, collocation can shed light on evaluative language (see Hunston, 2011). Essentially, collocation analysis helps establish “how words function in context” (Mahlberg, 2014, p. 220).

Generally, the word under study is referred to as the ‘node’ and a word it co-occurs with is labelled a ‘collocate’ (Sinclair et al., 1970/2004, p. 10). Some research in corpus linguistics distinguishes between the general concept of ‘co-occurrence’ (i.e. words appearing close to

each other in a given text) and the more specific concept of ‘collocation’ as quantified by the strength of association between node and collocates (see e.g. Evert, 2008). This convention is adapted in the present study. Evert (2008, pp. 1220–1221) points out that the “Firthian” approach to collocation is mainly based on ‘surface co-occurrence’, where the span of collocation is defined by proximity in word tokens. Other possible types of co-occurrence are ‘textual’, e.g. following sentence boundaries rather than word-based spans, or they are ‘syntactic’, based on grammatical patterns (Evert, 2008, pp. 1222–1223).

Many different types of statistical collocation association measures have been put forward to identify which word pairs are most strongly attracted to each other (for an overview of the most common ones see Evert, 2008, sec. 5). Reviewing the state of the art of collocation research, Gries (2013, p. 159) argues that “after many decades of ‘more of the same’ [...], it is time to explore new ways of studying collocations”. The new approaches that he suggests are (i) “directional measures” (that take into account word pairs that are not symmetrically attracted), (ii) “dispersion”, (iii) “type-token distributions and/or their entropies” (arguing that the counts in the contingency tables for conventional association measures are too simplistic), and (iv) “extendability to multi-word units” (Gries, 2013, p. 159).

One point that is not mentioned in this outlook is comparison. Although collocation is such an important concept in corpus linguistics and corpus linguistic work is said to be inherently comparative (see Section 2.3), in the past decades of collocation research little emphasis has been placed on creating comparative approaches. An early attempt at comparing texts based on collocation is presented in the OSTI report (Sinclair et al., 1970/2004). The researchers set out to use collocation association information to compare a corpus of spoken conversation and a corpus of popular scientific writing. A third “disputed” corpus also contained popular scientific writing but – for the sake of testing the method – was considered

to be unknown. The authors used a variety of complex statistical procedures to (i) test the collocational significance of a set of “test pairs” in the spoken and written corpus based on an association measure and (ii) to calculate likelihood ratios for the test pairs in the disputed text based on the association measures in the spoken and written corpora. The results showed that the method was able to correctly match the disputed text to the written corpus.

Despite the promising results, the method does not seem to have been taken up much by later studies, possibly because the report was only officially published in book form thirty years after it was originally completed. The conditions for the study were hindered by the drastically limited computer power compared to the technology available in the 21st century. Therefore, Sinclair et al.’s (1970/2004) comparison was restricted to word pairs occurring at least ten times in the text and a span of only one position to the right. Another potential reason for little uptake might be the indicated statistical complexity and the associated assumptions and approximations that are introduced into the comparison, as the authors themselves acknowledge (Sinclair et al., 1970/2004, p. 133), probably due to the technical limitations in handling corpus data at the time.

In recent years, initial interest has developed into the identification of diachronic collocates in order to detect discursive change over time. The *DiaCollo* tool was developed to assist historians in analysing discursive changes over time (Jurish, 2018), such as the evolving lists of proper nouns collocating with the German noun *Krise* (“crisis”) (Jurish, 2015) or changing hyponyms of the noun *Getränk* (“drink”) collocating with the verb *trinken* (“to drink”) (Jurish, Geyken, & Werneke, 2016). *DiaCollo* tool provides the collocation association measure for one or two nodes in the selected corpus “slice” (Jurish, 2018). As the user operates a slider to visually move through the corpus, the visualisations are updated according to the collocation association scores for a particular slice. The tool can also explicitly compare two

separate subcorpora; in this case the absolute difference between the association scores is used as the basis of the visualisation. Similarly, Kehoe and Gee (2009, p. 267) present a heat map visualisation that indicates changes in the z-score collocation association measure across monthly subcorpora.

It is not straightforward what exactly the implications are when individual lists of collocates – based on statistical association tests – are compared with each other for overlaps because of the various assumptions that have been made for each of the tests. Colleagues and I have put forward a method of ‘co-occurrence comparison’ between corpora that directly compares co-occurrence counts rather than collocation association measures (Wiegand, Hennessey, Tench, & Mahlberg, 2017b). The method is available in the *CorporaCoCo R* package (Hennessey et al., 2017) and has the advantage that the significance testing (using a Fisher’s Exact Test) happens at the point of comparison of co-occurrence patterns between the corpora. As all analysis chapters of this thesis make use of the *CorporaCoCo* method, it is explained in detail in Section 3.3. This procedure bears similarity with earlier proposals of ‘key collocates’ by individual researchers (Mahlberg & O’Donnell, 2008; Durrant, 2009), based on log likelihood comparisons. These proposals have not been taken up further or formalised in a tool.

A simpler comparative concept termed ‘consistent collocates’ has been used in the study of discourse in newspaper corpora (Gabrielatos & Baker, 2008); here, unlike in the OSTI comparison, the focus is on similarity. Accordingly, collocates are computed based on association measures for each annual subcorpus and those collocates that qualify as significant in the majority of the subcorpora (the exact definition depends on the study) are considered “consistent”. This idea of long-lasting collocates links with Gries’s (2013) emphasis of dispersion criteria for collocation.

In recent years, interest has been revived in the idea that

[c]ollocates of words do not occur in isolation, but are part of a complex network of semantic relationships which ultimately reveals their meaning and the semantic structure of the text or corpus. (Brezina, McEnery, & Wattam, 2015, p. 141)

Early work on collocation networks is found in Phillips (1985) and Williams (1998, 2002). Brezina et al. (2015) have developed the tool *GraphColl* (now part of the software *LancsBox*), which calculates collocates, offering a wide selection of statistical collocation measures, and then plots these in networks, starting from a central node. They propose an additional collocation criterion for Gries's (2013, p. 159) list, 'connectivity' (Brezina et al., 2015, p. 141), arguing that networks help understand this component of collocational meaning. The network representation opens up further interdisciplinary opportunities, such as applying concepts from graph theory (see P. Baker, 2016). My analysis in Chapter 4 builds on the concept of collocation networks and discusses their links to other corpus linguistic concepts.

2.4.5 Keyness and topicality

Several decades ago Raymond Williams (1976/2011, p. 13; emphasis in original) compiled a list of cultural keywords as "the record of an inquiry into a *vocabulary*: a shared body of words and meanings in our most general discussions, in English, of the practices and institutions which we group as culture and society". Although not a linguist himself, his approach of capturing the zeitgeist in cultural keywords (including such diverse words as *equality*, *native*, *mediation*) has had an impact on linguistics. For example, Stubbs (2001) illustrates how corpus methods can be used to shed light on how culturally salient words are used in language. Mahlberg (2007c) studies the cultural keyword *sustainable development* in newspaper articles from the year of an important climate summit, 2002, showing that there is much disagreement about its meaning

(or lack thereof). Teubert and Čermáková (2004) do not refer to cultural keywords specifically, but raise similar issues in their case study of *globalisation*. They find that its meaning continues to be negotiated in parts of the discourse community even when it has already settled in other parts (see Teubert & Čermáková, 2004, p. 140).

Scott (1997) introduced an empirical method of identifying keywords, based on frequency rather than theoretical reasons or intuition. This procedure is based on comparing a wordlist of the study corpus to that of a ‘reference corpus’ using a statistical significance test. Words that occur significantly more often in the study corpus than in the reference corpus are identified as “key”. Keyword analysis has proven to be a very popular method, particularly for discourse and stylistic analyses (see Mahlberg & Wiegand, 2018 for a review of corpus stylistic studies). Corpus linguists often start exploring the data by gaining an overview of a corpus via frequency or keyword lists in order to identify aspects of interest for detailed analysis. Regular wordlists simply show highly frequent grammatical words and very general content words at the top. By contrast, the top ranks of a keyword list shows what the text “boils down to” (Scott & Tribble, 2006, p. 56). The keyness, so the quality of being key, is associated with the ‘aboutness’ of a particular text (e.g. as content words and proper nouns) and can also indicate a certain style (Scott & Tribble, 2006, p. 55). Although keywords do not show discourses per se, they can point to “traces” of these (P. Baker, 2004a, p. 347). The examination of (a selection of) the keywords in context, for example in concordance lines or the generation of their collocates and ‘clusters’ – “repeated sequences of words” (Mahlberg, 2007a, p. 5) – can then point to salient discourses in the corpus (e.g. P. Baker, 2006).

Different perspectives on comparisons can be achieved by changing the reference corpus. In an article on news discourse related to “sleep and sleeping disorders”, Seale, Ziebland, and Charteris-Black (2006, p. 2577) follow an approach that they call ‘comparative keyword

analysis'. Their comparison is not based on one reference corpus, but several corpora under examination are cross-compared. Similarly, P. Baker (2004b, 2006) studies discourses in parliamentary debates on the gay male law reform and fox hunting via keyword lists generated by comparing the two sides of the debates and in fact encourages the use of multi-direction keyword comparisons in order to counter-balance the focus on difference. Another variation of the keyword method is a text-internal comparison. This is, for example, particularly suitable for stylistic studies comparing the dramatic speech contents and style among different characters in a play (see Culpeper, 2009).

An extension of the keyword to a 'key cluster' approach allows the researcher to compare phrases rather than only single words across corpora. Key clusters can be helpful for identifying 'local textual functions' of specific lexical items in a particular (set of) text(s) (Mahlberg, 2007a, p. 4), as illustrated in Mahlberg's (2007a) comparison of clusters in Dickens's novels and a reference corpus of 19th century fiction. Working on the level of texts rather than lexical items, Anthony and Baker (2015b, p. 273) argue that keyness can aid the selection of "prototypical texts for close reading". They have developed the *ProtAnt* tool (Anthony & Baker, 2015a), which ranks texts by the number of keywords they contain – with the assumption that texts with the most keywords in a corpus are prototypical of that corpus. The concept of keyness has further been transferred from the lexical to the semantic level in Rayson's (2003, 2008) online tool *Wmatrix* (<http://ucrel.lancs.ac.uk/wmatrix/>), which allows the semantic tagging of a corpus, and the subsequent generation of 'key semantic domains'. I explain the methodology of semantic tagging in Section 2.4.6.

As the popularity of keyword research has increased, the statistical methods of keyword identification and their implementations have received some criticism. Kilgarriff (2005) discusses several pitfalls of using models assuming that words are randomly distributed; this

has to be kept in mind when keyword analysis is carried out with tests making this assumption. Keyword analysis has also been linked to the temptation to focus only on differences when working with corpus data (see e.g. P. Baker, 2004a). Further supplementary analyses can be carried out to compensate for the focus on difference. These include to generate so-called ‘key keywords’, which are key across many texts in a corpus (Scott, 1997), and therefore provide a high level of aboutness of a corpus. More abstract meaning relationships can be explored via links between key keywords, which Scott and Tribble (2006) map as a network. I return to these networks and their similarity to collocation networks (see Section 2.4.4) in my analysis in Chapter 4. Key keywords have a relatively high dispersion by default, so that they can arguably provide insights into similarity – at least across subcorpora, analogous to the consistent collocates procedure (Gabrielatos & Baker, 2008; see Section 2.4.4). In an overview of keyness analysis, Gabrielatos (2018, p. 250) has recently introduced the notion of ‘keyness-S’ for a keyness measure indicating similarity, representing “the absence of difference”, in addition to ‘keyness-D’ for difference. Aside from improvements in the statistical procedure, P. Baker (2018, p. 78) reminds us that “researcher reflexivity” is crucial for honestly dealing with limitations of keyword analysis and developing this area further.

Topic modelling represents an alternative approach to aboutness that is not as widespread in corpus linguistics as keyword analysis, because it originates from machine learning. Unlike keyword analysis, topic modelling is based on co-occurrence data. Rather than highlight individual words, its output is a list of topics, each of which is formed of a set of words that occur in similar contexts. Murakami, Thompson, Hunston and Vajn (2017, p. 270) argue that there are similarities between topic modelling and collocation networks (see Section 2.4.4), due to their shared source of co-occurrence data. However, the authors suggest collocation networks

capture the “distributed or prosodic meaning associated with phraseology whereas topic models identify thematic meaning” (2017, pp. 270–271).

2.4.6 Semantic tag analysis

As discussed in the context of text-internal parameters for corpus compilation, lexis is the raw data of corpus analysis. Some approaches have been developed to go beyond the lexical level. However, as Mahlberg (2005, pp. 44–45) notes, the issue of annotating a corpus (e.g. with part of speech – POS – tags) is viewed as controversial by some. Sinclair (2004a, p. 48) distinguishes ‘mark-up’ that preserves information of the original text in a corpus (e.g. intonation in speech or italics in a written text) from ‘annotation’ that is based on analytical categories. In his view, annotation can skew the analysis of the actual text, especially if the tagger is based on a traditional model of grammar (i.e. recognising ‘nouns’, ‘verbs’, ‘adjectives’ etc.) that has not been developed based on corpus evidence. Nevertheless, a large body of corpus work has been established based on annotated corpora. For example, POS tags can be useful when searching for a particular sense of an ambiguous word form in a large corpus like the British National Corpus (see e.g. Weisser, 2016, Chapter 8), whereas a strictly Sinclairian approach would rely on distinguishing these senses via co-occurrence patterns (e.g. in concordance lines).

The use of POS tags for accessing particular senses already goes some way towards a more efficient, (semi-)automatic identification of word senses. To fully identify these senses and group them into semantic domains, however, another tagger is required that can identify additional information beyond the part of speech. The *UCREL Semantic Analysis System* (*USAS*) tagger provides this functionality for annotating individual words and a range of multi-word units in a corpus according to 21 “major discourse fields” and 232 subcategories (Archer, Wilson, & Rayson, 2002, p. 2). Each of the major discourse fields is labelled by a capital letter,

such as “M” for “Movement, location, travel & transport” and “S” for “Social actions, states & processes”. These semantic domains have been developed for content words; “[w]ords belonging to closed classes (such as prepositions, conjunctions, and pronouns), as well as proper nouns, are marked by a tag with an initial Z” (Rayson, 2003, p. 66). The *USAS* tagger first assigns candidate semantic tags to the POS-tagged word forms based on a lexicon that the developer team extends as it tags a wider variety of texts (see Rayson, 2003). In a second phase, the candidate tags are “disambiguated” based on a set of techniques, i.e. rearranged in the most likely order of appropriately describing the semantic domain of a given lexical item in its context. These techniques make use of various sources of information including the POS tags from the first phase, “general likelihood ranking” (i.e. which sense of a word form is generally considered more frequent) and the “domain of discourse” among others (Rayson, 2003, pp. 67–68).

The online interface *Wmatrix* (Rayson, 2003, 2008; <http://ucrel.lancs.ac.uk/wmatrix/>), allows users to upload their own corpora, to automatically carry out the POS and semantic tagging and, finally, to analyse the tagged corpus with common corpus linguistic methods such as keyness comparisons and concordances. Based on the semantic tags, *Wmatrix* can generate key semantic domains, i.e. semantic categories that are overused in one corpus compared to another, as mentioned in Section 2.4.5. This tool has been applied in various areas including discourse analysis (e.g. Cheng & Lam, 2013; Prentice, Rayson, & Taylor, 2012) and stylistics (e.g. Balossi, 2014; Mahlberg & McIntyre, 2011). Prentice et al. (2012) make the case for applying another traditional corpus linguistic method to semantically tagged data: collocation. They generate ‘semantic tag collocation’ by replacing the word tokens with tags when computing conventional collocation association measures:

[...] the node is a word, such as *son* and the collocate is a semantic category, such as Personal Names, or vice versa. Semantic tag collocations give a view of the terms most commonly associated with particular concepts or vice versa. (Prentice et al., 2012, p. 275)

Wmatrix (version 4, 2019) provides a beta collocation tool that allows the user to generate collocates as individual word forms for a given semantic tag and vice versa based on association measures.

2.5 Discourse domains

This thesis considers ‘public discourse’ from three different perspectives, based on three different types of data: academic discourse, digital discourse and news discourse. These different types of sources have inherently different register norms based on different social contexts and expectations. It is likely that they are written by different people; whereas the academic articles are written by experts in surveillance studies, the news articles are written by journalists who are also professional writers but may not focus on surveillance and operate in a different social context. The rationale behind sampling blog posts was to include voices from outside the authoritative institutions of academia and the media. Note that it is still possible for the writers to overlap: for example, the Surveillance Blog Corpus that I analyse in Chapter 5 in fact contains posts from the editor of the *Surveillance & Society* academic journal analysed in Chapter 4. However, the act of posting a piece of writing to a blog (especially if it is a personal rather than institutional blog) implies a different social context from publishing an official academic article in a journal. The following subsections introduce the three domains studied in this thesis: academic discourse (2.5.1), digital discourse (2.5.2) and news discourse (2.5.3).

2.5.1 Academic discourse

One of the aims of the present study is to compare the representation of the concept of surveillance in media discourses with academic discourses (see Chapter 4). In this section, I therefore provide a brief summary of relevant research on academic discourse.

In many cases the label ‘academic discourse’ is used to refer to written rather than spoken language, although some research has also been carried out on academic presentations (Hood & Forey, 2005; Hyland, 2009) or lectures (e.g. Alsop, 2016; Flowerdew, 1991, 1992). So, studies of academic discourse are often concerned with investigating general features of the academic writing style, sometimes by comparing various disciplines (see Triki, 2019) or looking specifically at the linguistic features of interdisciplinary academic reports (e.g. Murakami et al., 2017; Teich & Fankhauser, 2009; Teich & Holtz, 2009). Indeed, a study using multidimensional analysis has shown “that some of the variability within registers of academic writing can be accounted for by publication type and discipline variation” (Egbert, 2015, p. 26). When specific disciplines and/or particular topics are chosen, the focus often lies on language features such as showing the importance of formulaic language in contributing to discipline conventions (e.g. Bondi, 2009; Pecorari, 2009). Some research specifically focuses on learner writing. Similarly, a lot of research on published texts aims to analyse the conventions of successful writing in order to improve teaching methods in academic writing. For a recent review of research in this area, see Chen and Flowerdew (2018).

A different and less common approach to academic discourse is more concerned with the aboutness of the discourse in the field rather than the register style. Examples of this type of research include studies with a clearly non-linguistic focus, like a comparison of climate change representations across the domains of academic, political and media discourses in Germany by Weingart, Engels and Pansegrau (2000). Based on their findings, Weingart et al. (2000) argue

that the different perspectives and communicative purposes of the three domains lead to divergent communication styles and bear the risk of misunderstanding. The study shows how discursive representation, particularly of a controversial topic, can differ between various domains and highlights the dangers related to a divide between academic and media discourse.

With regard to surveillance discourse, Mehrabov (2015) has carried out a content analysis of research articles published in the journal *Surveillance & Society* (S&S) in order to determine the main themes. These themes guide the co-occurrence analysis of trends across the S&S Corpus in Chapter 4. Another study of research articles that my analysis (see Section 4.2.2) is partly based on is Taylor's (2008) study of articles in corpus linguistics. Taylor (2008) uses corpus linguistic methods in a self-reflective manner to examine how researchers in her own area describe the nature of corpus linguistics as a method or discipline. She applies concordance, collocation and keyword/cluster analyses to the term *corpus linguistics* in academic articles. The study finds "radical differences in the representation and understanding of what corpus linguistics is" (Taylor, 2008, p. 196). At the same time, Taylor's (2008) approach seems rather unique in corpus linguistics with its emphasis on examining the topic of academic discourse with an interest in corpus linguistics itself.

As part of my analysis of the academic domain of surveillance discourse, I study the range of definitions of *surveillance*. Flowerdew's (1991, 1992) study of science lectures is relevant in this regard, because he analyses the speech act of defining. There is limited recent research on definitions, with the exception of Triki (2019), who compares definitions in research articles from linguistics and computer science. A lot of the early work on definitions, like Flowerdew's research (1991, 1992), appears to have been motivated by pedagogical concerns (see the overview in Pearson, 1998). Academic definitions were mainly analysed in order to teach students how to understand and produce their own definitions (e.g. Selinker, Trimble, &

Trimble, 1976; Swales, 1984). Flowerdew's (1991, 1992) research on definitions in science lectures is one of the first systematic corpus-based approaches to definitions. Pearson (1998) then developed a method of semi-automatically extracted definition formulae from corpora. Her corpus findings show that definitions have performative functions and that context is an important factor.

2.5.2 Digital discourse

Activities of daily life are increasingly carried out digitally. This involves not only communication (instant messaging, e-mail, etc.) but also actions at the interface of the virtual and the material (see e.g. Jones, Chik, & Hafner, 2015), such as watching a film via a streaming service like Netflix and ordering books or groceries online. A large body of research on 'computer-mediated communication' has developed over the last two decades; an overview of the early work is provided in Herring (2004). Since then, the term 'discourse' has been used to refer to research in this area, for example in the edited volume *Digital Discourse: Language in the New Media* (Thurlow & Mroczek, 2011a) and a recent handbook chapter on "Computer-mediated discourse 2.0" (Herring & Androutsopoulos, 2015). Thurlow and Mroczek's volume focuses on discourse as language in use, as they emphasise in the introduction (see Thurlow & Mroczek, 2011b, p. xxiii). Referring to Androutsopoulos's (e.g. 2010) work, they highlight the potential of digital discourse analysis to

move beyond a one-track interest in the formal features of new media language (e.g., spelling and orthography) and a preoccupation with delineating individual discourse genres; instead, greater attention should be paid to the *situated* practices of new media users (Thurlow & Mroczek, 2011b, p. xxi; emphasis in original)

By recognising users' interactions with digital media as 'situated practices', researchers can account for the implications that these media have in social life rather than assuming that they are one-to-one transformations of analogue to digital formats. For example, Blood (2002, p. 9)

calls the weblog “a form that is native to the Web” that was not simply transferred from any analogue format (e.g. a diary or scrapbook), but is instead built on the principle of “continual publishing” that is only possible digitally. The forms of social interactions enabled by the “interactive writing spaces” of the digital media differ considerably from traditional, analogue forms of conversation and writing, to the point that “[t]hese new practices are also challenging the ways discourse analysts think about texts, social interactions, and even the nature of language itself” (Jones et al., 2015, p. 1).

One of the practices that has been highlighted as particularly important in the context of interactions via digital media is intertextuality. Intertextuality is an attribute of discourse in general (Jones et al., 2015; Teubert, 2010). Yet, the “technological affordances” of digital discourse offer new and easier ways of creating intertextual links (Jones et al., 2015, p. 6).

The digital discourse analysed in this thesis is one of the interactive, web native text types: the blog post. Blog research is undertaken in many disciplines from different perspectives. The quote on digital literacies highlights how online activity matters and blogs appear to hold a special place in the development of the interactive web. ‘Blogs’, short for weblogs (see e.g. Meinel, Berger, Bross, & Hennig, 2015), emerged in the late 1990s as special websites run by individuals. However, compared to the relatively stable format of a personal homepage, a weblog tended to be dynamic due to regular updates, often on a daily basis. In one of the earliest and widely cited practical introductions, the *Weblog Handbook*, Rebecca Blood provides an insider’s perspective that does not only describe the development of weblogs but also acts as a guide to good practice. She explains the development and function of the early weblogs as follows:

“Links with commentary, with the new stuff on the top” was the formula; for those who found them, these sites served as a welcome guide through the increasingly complex World Wide Web. (Blood, 2002, p. 3)

Whereas at first the creation of a blog required advanced technical skills, blogging became more widespread as web services were developed that made it easy for the general public to set up, design and maintain their own blogs (for detailed accounts of the history of (we)blogs, see e.g. Blood, 2002; Rettberg, 2008). This creates a stark contrast between the blog posts and the two other presumably much more edited text types analysed in this thesis (academic articles and newspaper articles). In particular, the entrance barrier to publishing a research article is higher than that for publishing a blog post, allowing the blogger to react to current affairs much faster than would be possible for an academic publication. Many academics have actually created blogs as additional platforms for presenting and developing their work. Zou and Hyland (2019) point out that academic blog posts have a much more heterogeneous audience than research articles. As a result, they find that academics use a more personal and evaluative style in blog posts in order to highlight the value of their research to a wider audience (Zou & Hyland, 2019).

Hyperlinks played a central role in the earliest blogs which were often used to literally “log” the maintainer’s journey through the internet (see Blood, 2002, pp. 3–4). As blogs developed further, hyperlinks still tended to serve important functions and link blogs to many other sources such as other blogs, Wikipedia, Amazon or news websites (Myers, 2010, Chapter 3). Importantly, they represent digital affordances for intertextuality (see Jones et al., 2015).

Online platforms are in constant development. So, compared to the initial excitement about blogging technology, “[e]nthusiasm for and popular media coverage of blogging as a distinct publishing format in its own right has declined considerably since its heyday in the early 2000s” (Bruns, 2017; online). At the same time, Bruns (2017; online) points out that features from blogging platforms have been incorporated into social media platforms, which are often called “micro-blogging platforms”.

Overall, online platforms are an important source for analysts interested in ‘public discourse’. Teubert (2012, p. 113) goes as far as to say that “[t]he internet has indeed changed our understanding of public discourse” due to the debates that it makes possible. An example of how blogs can be part of the political discourse is given by Bakir (2010). She focuses on the role of web users’ interferences with strategic political communication in the 2003 Iraq War. One of Bakir’s (2010, Chapter 2) case studies focuses on the blog of Salam Pax,² who was one of the first Iraqi bloggers posting in English and therefore providing a local perspective on the Iraq war to a world-wide, and particularly Western, audience. Bakir’s (2010) discussion of the blog centres on issues of authenticity: at a time when doubts had been raised concerning the trustworthiness of Western politicians and media in relation to the Iraq war, Pax’s blog provided a refreshing and personal perspective. This blog appears to be an example of individual blogs that “have raised the stakes of local coverage, because they have been able to converge the consumption of local information with the production of local information” (Gordon & de Souza e Silva, 2011, p. 119) where local newspapers have failed to do so.

Some commentators are more sceptical about the chances of blogs and similar venues to allow people to actively shape public discourse. Teubert’s (2012) analysis focuses on the opportunities that comments on news blogs provide and the role this medium can play in democracy. Although he finds that stimulating debates are happening, Teubert (2012) expresses doubts that the ideas developed on these platforms will have any wider impact on society. He argues that the different groups are now characterised by a division between “us” and “them” rather than a sense of solidarity (Teubert, 2012, p. 120).

² The blog was active until 2009, but is still available online at the time of writing (see Pax, n.d.).

Among corpus linguistic studies of blogs, a continuum of qualitative to quantitative research focuses can be identified. Located at the more qualitative end are studies like Hoffmann's (2012) analysis of cohesion in personal blogs and Puschmann's (2010) detailed account of corporate blogs. The quantitative method of multi-dimensional analysis has been used to identify register differences within blogs (Grieve, Biber, Friginal, & Nekrasova, 2010) or between blogs and other web registers (Titak & Roberson, 2013). Biber and Egbert (2016) similarly compare web registers, taking a more bottom-up approach to register boundaries: their data is based on crowd-sourced register divisions rather than a-priori definitions of what a blog is. To give a final example, Elgesem et al. (2016) use a combination of methods, including the quantitative approach of topic modelling and keyword analysis, to study blogs covering the story of Edward Snowden's leak of classified documents (cf. Section 2.2; also see Elgesem & Salway, 2015).

Obviously, blogs are not the only online medium that has been studied linguistically. The "microblogging" platform Twitter has received much attention from linguists. Because of the large quantity of tweets posted on the platform around the clock, it is possible to quickly collect a large social media corpus from Twitter. For example, the platform has been used to study the emergence of neologisms and slang (Grieve, Nini, & Guo, 2017). However, the brevity of Twitter posts restricts the linguistic methods that can be used for analysis (though see the categorical form of multidimensional register analysis developed for Twitter by Clarke & Grieve, 2017).

In the surveillance studies literature, blogs have also been recognised as relevant sources for public discussions of surveillance. The following quote from a period in which blogs had just recently become popular illustrates this point:

While it would be a mistake to ignore the contribution of popular culture to understanding surveillance, there are decided limits to what can be said. Much work remains to be done

in exploring the connections, some of which may turn out to be important in ways that we cannot guess at today. The growth rate of new systems for blogging and interactive sharing of ideas and images on the internet [...] alone means that this field of study is likely to be a growth area. (Lyon, 2007, p. 158)

Lupton (2015) describes a mutual relationship between the traditional and social media. Not only do social media users quickly circulate links to online news articles (and so contribute to an article's visibility), but Lupton (2015, p. 161) argues that journalists increasingly rely on information from social media posts, with the potential to "perpetuate the rumours in their own tweets and online 'breaking news' stories". Social media are also an increasingly important domain to consider for surveillance scholars. For example, abundant sharing of a post describing an individual's bad behaviour could exaggerate "minor wrongdoings" (Lupton, 2015, p. 161) and can leave a permanent online trace. In some sense we have a right to opt out of social media, although this may be at the risk of becoming marginalised in societies with a strong emphasis on the digital (see Bauman & Lyon, 2013, p. 29). On the other hand, it becomes increasingly difficult even for people who do not wish to be mentioned on social media or shown in photos to avoid this, as so many others around them will be active users (Lupton, 2015). Another reason that new media trends are important to consider in both media and surveillance studies is that media providers, in collaboration with advertisers, progressively make an effort to create a customised experience for readers (Turow, 2011).

2.5.3 News discourse

Newspapers are a popular choice of data for corpus linguistic studies of discourse. Reasons for this popularity include easy access to large amounts of data and the power that newspapers are thought to hold in disseminating ideologies as well as a commenting on current events. Newspapers offer a platform for the "expression of the dominant values in society while

allowing powerful new forms of social identification through those values” (Conboy, 2010, p. 81).

Over the past years, the media landscape has changed radically. The newspaper industry has been facing many challenges, regarding both sales and advertising revenue, and the need to set up a profitable online presence (Leurdijk, Nieuwenhuis, & Poel, 2014). While few newspapers have been economically successful at this, a complete change to online publishing does not seem feasible either (Leurdijk et al., 2014). Newspapers that move to an online-only format may gain readers, but the time that their audience spends reading their content can fall drastically, as Thurman and Fletcher (2018) show for the British *Independent*. Yet, Conboy (2010) reminds us that the history of newspapers testifies that the medium has previously survived many technological advances.

In traditional media studies, much work has relied on the concept of ‘news values’. This is usually attributed to Galtung and Ruge (1965) and has been further developed by Bell (1991, p. 155), who argues that “[t]he values of news drive the way in which news is presented”. Bell’s (1991, pp. 155–160) adaptation contains three groups of news values: (i) “[v]alues in news actors and events” including “negativity”, “recency”, etc.; (ii) “[v]alues in the news process” such as “continuity” and “competition”; and, (iii) “[v]alues in the news text” comprising the three values of “clarity”, “brevity” and “colour”. As Bell (1991) acknowledges, the functions of these groups overlap to some extent: for instance, the value of “superlativeness” of an event – from group (i) – will naturally also be reflected in the language use. Perhaps it is possible to compare this conceptualisation with the corpus linguistic notion of keywords (introduced in 2.4.5), which describe both a text’s style and contents. In some ways, Bell’s (1991) first category represents the aboutness of a text in terms of both relevant topics and proper names, while the last category can be understood as more of a stylistic indicator.

Bednarek and Caple (2014) argue that the concept of news values is not sufficiently embedded into linguistic analysis of news discourse, suggesting that unfamiliarity with the concept may be the reason for this gap. In order to encourage more empirical linguistic investigation in this area, they provide a set of practical guidelines to identify news values in a corpus. These are based on the analysis of frequency and dispersion information as well as other standard corpus linguistic measures. Further qualitative, context-based investigation is then carried out based on the statistics. A study of news values in the coverage of Hurricane Katrina in the US demonstrates how this approach can be usefully applied to the analysis of a news event in a specialised corpus (Potts, Bednarek, & Caple, 2015).

Corpus linguistic analysis of media discourse requires the data to be in a suitable electronic format. While this can be complicated for spoken news genres, it is nowadays relatively simple to achieve for written data from sources like newspapers (O’Keeffe, 2012). As many newspapers are now available in electronic format, newspaper articles have become a popular source of data in corpus linguistics. Previous studies have mostly investigated the newspaper representation of particular groups of people, such as refugees and asylum seekers (Gabrielatos & Baker, 2008, p. 5; also see Partington et al., 2013) and trans people (P. Baker, 2014), or a particular topic like sustainable development (Mahlberg, 2007c) and climate change (Bevitori, 2010; Grundmann & Scott, 2014). Another line of research is more concerned with the nature of “newspaper language” itself, as in the edited volume *Exploring Newspaper Language* (Andersen, 2012), for instance looking at neologisms arising from particular news events (De Smedt, 2012).

A popular method of accessing newspaper articles for the compilation of newspaper corpora (for instance used in P. Baker, Gabrielatos, & McEnery, 2013; Gabrielatos & Baker, 2008; Grundmann & Scott, 2014; Partington et al., 2013) especially with regard to UK data, is

to download them from a newspaper database such as the Nexis News Search (LexisNexis, 2019). Disadvantages of relying on Nexis UK include that only 500 articles can be downloaded at once and that researchers can neither be certain whether the dataset contains all articles from the sources nor exclude the possibility of duplicate articles. These limitations can perhaps only be overcome (or at least controlled) with a customised programme, as used by Grundmann and Scott (2014) for cleaning their Nexis UK corpus. Some projects, such as the Norwegian Newspaper Corpus (Andersen & Hofland, 2012, p. 9), are based on directly “crawling” newspaper websites. While this allows the researcher more control over the data collection, it also poses more technical challenges depending on the desired features of the corpus.

Just like it is impossible to study the discourse at large (see Section 2.3; Teubert, 2010), studies of newspaper discourse have to be selective. One method of selection is the number of newspaper sales. Another option is to select the publications based on readership estimation figures, as other factors like the availability of free online articles and shared newspapers mean that more people may be exposed to the content than sales figures suggest (P. Baker et al., 2013). While Teubert (2007b) argues that estimates are not reliable and that a measure of intertextuality would be most useful in order to assess the potential impact of a news story, this approach is less practical. Another criterion that the corpus compilation of UK newspapers is often based on is the distinction between tabloids and broadsheets (e.g. P. Baker et al., 2013; Partington et al., 2013; Seale, Boden, Williams, Lowe, & Steinberg, 2007), although drawing the line between the categories is not always a straightforward task (P. Baker et al., 2013).

Regardless of the source of newspaper articles, technical challenges arise while converting the original newspaper text to a format that is compatible with corpus tools, while not skewing the original data. For instance, technical errors can lead to duplicate articles (Andersen & Hofland, 2012). Moreover, it has to be considered that ordinary monomodal

corpora only contain texts that have been removed from their original context. As information on time, location and speakers in the case of spoken data or the visual presentation of written texts is missing, monomodal corpora can be subject to the criticism of potential decontextualisation (P. Baker, 2006). Some context can be embedded into the corpus in the form of mark-up. It is more difficult to compensate for the lack of visual material accompanying the original texts. Databases like Nexis UK only provide the raw text of the articles. A possible solution is to incorporate a qualitative stage into the project, during which a small sample of the corpus is chosen for further analysis by means of accessing scanned images of the newspaper pages, if available, or the original article on the newspaper website. Bednarek and Caple (2017, p. 8) make a strong case for what they call ‘corpus-assisted multimodal discourse analysis’. They argue that news photographs in particular “play a very important role in news storytelling”, but also point out that “other semiotic resources such as typography, layout, framing and colour” contribute to “the construction of news values” (Bednarek & Caple, 2017, p. 107).

For the analysis of newspaper texts, it can be useful to distinguish between different newspaper sections. Mahlberg (2007c) shows that there is a link between an item’s meaning and the section where it occurs. For example, *sustainable development* has a positive connotation in obituaries compared to a more sarcastic usage in feature articles of *The Guardian*. Kehoe and Gee (2009) follow a similar approach for the term *credit crunch*, also in *The Guardian*. They trace the diachronic development of the compound over the period 2007–2008 as it is gradually mentioned in wider range of sections.

In an analysis of *The Times*’s representation of the suffragette movement, Gupta (2015) finds news reports of the movement placed together with unrelated, but negative news. While the practice of combining several news reports into one article was common at the time, Gupta

(2015) makes a convincing case that the editorial decision to combine certain articles adds an additional layer of meaning. Gupta (2015, p. 110) terms this practice ‘suggestive placement’. O’Donnell et al. (2012) build on the concept of ‘textual colligation’ by running keyword comparisons across sections of newspaper articles. They identify text-initial keywords in order to analyse ‘nucleus patterns’ in news stories, i.e. patterns of words that are found at the beginning of news articles. Another way of looking at the textual location is to examine co-occurrence patterns not of individual words within articles, but of types of texts that are placed together.

A section that seems to have received relatively little attention from corpus linguists is classified advertising (i.e. short adverts of the width of one newspaper column). A potential reason for this decline in interest is that classified adverts might play a less important role in contemporary newspapers because of online advertising. However, this text used to have important functions in everyday life in the 20th century (Bruthiaux, 1996). For a historical perspective, see Görlach (2002). As part of my analysis in Chapter 6, I examine some British classified ads related to surveillance discourse from the 1980s to 2000s.

For my analysis of surveillance discourse in *The Times* in Chapter 6, I use a subset of The Times Digital Archive (TDA). The TDA includes all articles from the *Times* newspaper print edition between 1785 and 2008. The advantage of having the TDA as opposed to downloading articles from a database like *Nexis UK* is that the corpus is not limited to a researcher-defined sub-set, based on the query terms needed for databases. *The Times* (London), was established in 1785 and has had an important status in British society throughout history, as a “paradigm of political influence”, particularly with regard to its early period (Conboy, 2010, p. 85). P. Baker et al. (2013a, p. 231) suggest that *The Times* is “a fair proxy for general newspaper English”. Moreover, the circulation of *The Times* actually rose from the 1980s to the 2000s (see

Franklin, 2008, pp. 7–8), the period under study in Chapter 6 of this thesis. The TDA has been used, for example, to study the language of 19th century advertisements and changes in speech presentation over time (Jucker & Berger, 2014). Details on how the TDA data was processed for the present study are provided in Section 3.2.3.

2.6 Discourse coordinates

This thesis sees place and time as fundamental coordinates of discourse. As examples of the text-external parameters (Section 2.4.2.1), they are not only important coordinates for compiling corpora, but also for the analysis. Scollon and Scollon (2003, p. 21) point out that “[w]e live and act, we speak and we write in a world of real spaces and of real time”. These parameters are also crucial for conceptualising surveillance. Lyon (2007, p. 16) argues that “[c]oordinates are key. Anyone who can pinpoint the time and place of some event or activity already has a handle on the situation”. In the following, I first highlight studies dealing with discourse and place (2.6.1) and then outline examples of discourse studies focusing on the parameter of time. Both sections are necessarily selective.

2.6.1 Discourse and place

The importance of place has been recognised in what Jaworski and Thurlow (2010, p. 12) describe as the “‘spatial turn’ in the social sciences”. This section reviews important work on language and place, ending with the framework of surveillant landscapes that guides my analysis of the representation of place in the surveillance discourse of blogs (Chapter 5).

Due to the increasing interest in place across the social sciences and humanities, much work on language and place has interdisciplinary influences as can be seen in areas like linguistic landscape studies (see Barni & Bagna, 2015) and urban linguistics (Busse, 2019).

Corpus linguistic research on spatial information has been closely connected to the use of ‘Geographical Information Systems’ (GIS), which allow the visual mapping of place names mentioned in a corpus (e.g. Gregory & Hardie, 2011). This technique can then be used to interpret the regional variation of social meanings associated with a particular concept: for example, H. Baker et al. (2019) use ‘geo-parsing’ in combination with concordance analysis to study the geography of prostitution in historical texts from the 17th century.

Place is of particular importance for sociolinguistic research, for example in relation to links between linguistic variation and the place of origin and residence. As Myers (2006) points out, the question of origin can be complex and often requires a process of identity negotiation. A focus on identity in relation to place is also picked up in some corpus linguistic research. Busse (2019) discusses keywords from a corpus of interviews with Brooklynites to capture patterns of place-making in neighbourhoods of Brooklyn, New York City.

In his study of blogs, Myers (2010) suggests that this text type has an unusual relationship to place: “[b]logs, set in a blogosphere of other blogs, are placeless by default; they have to do something to signal place or we don’t think about it” (2010, p. 48). Whilst the blogger is set in a physical environment, this environment is not shared with the audience or even with the server that hosts the blog. In Myers’s (2010) sample, references to place are rather rare, supporting his point that this lack of place references makes those indications of place that do exist stand out. These are categorised into references to: (i) providing any information on the “About” page that says where the blog is based either physically or simply the URL; (ii) mentions of a particular place in the posts; (iii) including images of flags; (iv) indirect references to place via deictic expressions (more frequent than explicit references); (v) inclusive or exclusive uses of *we*; (vi) code-switching as a marker of “affiliation” rather than actual physical location; (vii) photographs (Myers, 2010, pp. 51–57). Despite the general ‘placelessness’ of the blogs, Myers

(2010, p. 50) observes that “[t]he blogosphere may be imagined as separate from the geographical world, but bloggers use the language of space to construct it”. For example, bloggers and commenters use spatial deixis like *here* or *there* to point to the each other’s locations. Similarly, Turnbull (2013, p. 315) notes in her analysis of a corporate website that “[s]pace in computer mediated communication can be described as virtual and unbounded, but at the same time we talk about websites as containers, as the expression ‘in the website’ clearly shows”.

On the other hand, digital discourse can be considered to be much part of “place-making” (see Busse, 2019) procedures. For example, Lyons (2018, Chapter 4), in her study of a San Francisco district, argues that Instagram posts published from and geo-tagged for that particular area inherently contribute to shaping the linguistic landscape. Lyons (2018, p. 83) introduces another term for the patterns found in these online contributions: ‘filtered landscapes’. Her analysis is concerned with the history and attitudes towards a particular place, with concerns such as language choice in the local linguistic resources such as shop names and billboards and economic issues like gentrification.

The recently developed framework of ‘surveillant landscapes’ offers a linguistic perspective on surveillance and space, describing environments that “read” and “write” their passers-by through architectural and technological features (Jones, 2017). Jones’s approach is grounded in mediated discourse analysis and develops the notion of ‘linguistic landscapes’. This notion was originally coined in a seminal paper by Landry and Bourhis (1997) and has since attracted much attention to the point that a field has started to form, as apparent from various edited volumes and the emergence of the journal *Linguistic Landscapes*. A linguistic landscape refers to language resources in a physical space such as billboards and other signs (e.g. see Lou, 2014 for a discussion of billboard advertisements in Chinatown, Washington,

DC). Linguistic landscape research increasingly acknowledges the role of people in “authoring” (see Malinowski, 2009) the (mainly urban) environments that they inhabit. Whereas initial studies tended to focus on quantifying the display of languages (or language varieties), the scope of the research has since broadened out towards more interdisciplinary combination of approaches and methods of analysis and more qualitative interpretations (Barni & Bagna, 2015).

Jones’s (2017) surveillant landscape framework recognises the interactive authorship of passers-by, mobile technology, and the landscape or built environment. He approaches surveillant landscapes from a framework with three perspectives that are derived from mediated discourse analysis (introduced in Section 2.3). The first perspective refers to the messages communicated at the surveillant site. Looking at a sign announcing CCTV surveillance, the analyst would record what information the operator provides about the activity and what aspects are emphasised via the syntax, (e.g. thematization) and design (e.g. font size). This perspective is labelled ‘discourses in place’ (Jones, 2017, p. 154). The terminology for the second perspective, ‘interaction orders’ (Jones, 2017, p. 169), originates from Goffman’s (1983) work. In the surveillant landscape framework, this describes the interpersonal and power relations that are enacted between the watcher and the watched. Finally, the perspective of ‘historical bodies’ covers two consequences following the act of watching in the surveillant landscape (see Jones, 2017, p. 178). First, this includes the effect that the semiotic material has on the watched subject; either at the moment of watching or for forming a longer-term habit to comply with the messages from the watchers. Secondly, the act of watching can leave an effect on the surveillant landscape in the form of data being recorded which itself may affect future actions and/or opportunities of the watched subject. This is the case at border checkpoints where we show our biometric passports, but also happens in less official contexts, as when we “check into” a place

on social media. Both examples in return can affect the person being watched or and can lead to

[...] linking behaviors observed at one particular time and in one particular place to countless other times and places, so that it is not just a matter of us carrying around the habit of being watched within our consciousness and bodily hexis, but also of our history of being watched literally following us from place to place like a shadow. (Jones, 2017, p. 180)

According to Jones (2017), this aspect of surveillant landscapes makes them an important subject of sociolinguistic research. Indeed, he argues that most linguistic landscapes can be considered surveillant landscapes as particularly with spreading mobile technologies, passers-by become more and more “legible” to the surrounding area, especially in cities. Jones (2017, p. 182) further links the surveillant landscape to the notion of surveillant assemblages (see Section 2.2), characterising them as “complex assemblages of discourses, bodies, technologies, and social relationships which help to regulate the flows of people, goods, and information through our societies”.

2.6.2 Discourse and time

The temporal dimension of discourse has been indicated at various points in this chapter. Meaning has been introduced as a concept that is prone to change – and change implies a time scale. Time, like place, is an important concept for both discourse and the concept of surveillance. This section focuses on the discourse coordinate of time from a corpus linguistic perspective, in preparation of my diachronic analysis of the surveillance discourse of *The Times* (Chapter 6).

Diachronic corpus linguistics has developed extensively in recent years with advances in both methodologies and new, large corpora specifically built for diachronic analysis. A multitude of diachronic approaches exist; despite their differences in motivation and

methodologies, they share an interest in linguistic developments over time. One of the aspects which can define the type of analysis is the focus period; whether the study is concerned with “recent language change” (Mair, 2006) or earlier historical periods. This focus period can then become one of the text-external criteria for compiling a corpus (see Section 2.4.2.1). For example, time can be one criterion for creating so-called ‘snapshot corpora’ – a set of corpora compiled of comparable sources from different periods – to then facilitate a diachronic comparison (McEnery & Hardie, 2012, p. 9).

A different approach is to work with one corpus that spans a longer time period and therefore needs to be “segmented” into temporal subcorpora. Marchi (2018, pp. 179–180) identifies three types of segmentation: ‘[t]ext-lifecycle segmentation’ (according to the units of the text, such as daily newspaper issues), ‘[t]op-down segmentation’ (deciding on a conventional unit such as a calendar year) and ‘[b]ottom-up segmentation’ (e.g. based on information of distributions within the data). Based on the research question for a specific project, the appropriate type of segmentation is applied. An example of the top-down approach is the compilation of the SiBol corpora, which contain the full text of three British broadsheet newspapers (Marchi, 2018, p. 178) from the years 1993, 2005 and 2013, and therefore act as snapshot corpora that can be compared. Note that the year is not always a top-down unit: for example, Molino (2019) focuses on annual reports. As these appear once a year, the year is the natural text-lifecycle unit. An example of bottom-up segmentation is given in Gries and Hilpert (2008), who use the technique of ‘variability-based neighbour clustering’ to identify stages in a diachronic corpus.

The choice of segmentation depends both on the characteristics of the data and the aims of the analysis. In general, the more sampling points in a given time frame, the higher the ‘granularity’ of a diachronic study will be (Gabrielatos, McEnery, Diggle, & Baker, 2012; also

see Davies, 2010, p. 448). Gabrielatos et al. (2012, p. 153) put forward a simple formula for determining the granularity of a diachronic corpus so that corpora can be compared in this respect: the number of sampling points divided by the time period of the corpus based on a temporal unit like years. Their analysis applies a “wave, peak and trough” method to identify statistically significant peaks (Gabrielatos et al., 2012, p. 165), which represent significant increases from a preceding “trough”. Another way to gain an overview of diachronic trends is to track developments in the relative frequency of linguistic features, as for example used by Zinn and McDonald (2017) in their analysis of risk discourse in *The New York Times*.

With access to diachronic corpus data – either in the form of multiple snapshot corpora or one diachronically segmented corpus – it is possible to use any of the corpus methods introduced in 2.4.3–6 for diachronic analysis. The diachronic approach in the present work is applied to co-occurrence (in Chapter 6), in a similar spirit to Krishnamurthy’s (2018, pp. 58–59) argument that “[m]eaning arises from context, and context is continually changing, so collocational changes are actually the process of meaning formation, and language change merely the sum of the collocational changes”.

Diachronic analysis is fundamentally a form of comparison. Section 2.4.4 noted that, so far, research on comparing collocation across corpora is limited. “Consistent” collocates (Gabrielatos & Baker, 2008) offer a method for tracking collocates that stay salient across a given number of subcorpora, in a similar way to key keywords (see Section 2.4.5). In recent years, some research has started to track collocational change over time (Jurish, 2018; Kehoe & Gee, 2009; see Section 2.4.4), which has largely relied on computing collocation association measures for individual subcorpora. Chapter 6 uses the novel method of co-occurrence comparisons (Hennessey et al., 2017; Wiegand et al., 2017b) to track collocational changes

through pairwise statistical comparisons of the raw co-occurrence counts. The underlying comparison method is explained in detail in Section 3.3.

2.7 Principles of meaning-making

This chapter has outlined a multitude of approaches to meaning, discourse, corpus linguistic methods and surveillance. On the basis of the reviewed literature and in order to analyse meaning-making patterns in surveillance discourse, this study puts forward the following three principles of meaning-making:

(i) Meaning evolves with the discourse

Because “language is a social phenomenon” (Mahlberg, 2005, p. 188; Teubert, 2001, p. 129), meaning is created when language is used in daily life, whether in speech or in writing. This view of discourse overlaps with what Mills (2004) describes as the “mainstream linguistic view”; it agrees with the basic definition of discourse by Brown and Yule (see 1983, p. 1) of discourse analysis as “the analysis of language in use”. The approach taken here literally considers a word’s meaning to be made up of all instances of that word in discourse (see Teubert, 2010). Accordingly, meaning can never be fully defined, because the analyst – and the language user – is only able to obtain a tiny fraction of the discourse for analysis. In order to analyse language in use, and from various perspectives, all analysis chapters of this thesis deal with specifically compiled corpora from different social contexts.

(ii) Meaning emerges via comparison

From a corpus linguistic view, we can try to capture meaning by comparing patterns across corpora (see e.g. Koteyko, 2006). This also resonates with Teubert’s (e.g. 2001, 2005b, 2010)

view that discourse consists of meaning negotiation; through comparison and contrast, patterns of intertextuality and evaluation can become obvious. The approach of this thesis is inherently comparative in that it looks at three corpora from very different discourse domains and situational characteristics, that, nevertheless all relate to surveillance.

(iii) Meaning takes shape in co-occurrence patterns

As Section 2.4.4 on collocation and co-occurrences has demonstrated, this is a fundamental concept for any corpus linguistic work. All analysis chapters in this thesis rely on identifying co-occurrence patterns – not only raw co-occurrences, but also comparative co-occurrence results, in line with principle (ii). This work is based on the output from different stages of collaborative research for the development of the *CorporaCoCo* co-occurrence comparison package (Hennessey et al., 2017) that I have contributed to.

3 Methodology: Corpora and analysis frameworks

3.1 Introduction

Based on the three principles of analysing meaning introduced in Chapter 2 – (i) as a product of language use; (ii) identified via comparison and (iii) signalled by co-occurrence patterns in discourse – this thesis establishes a corpus linguistic methodological framework of co-occurrence comparisons for the study of surveillance discourses. This chapter presents the corpora and the methodology of the analysis followed in this thesis.

The main research question of this thesis is “how is surveillance discursively represented?”. The analysis is broken down into three chapters (Chapters 4–6), each of which focuses on one domain of public discourse. Given the different data sources, the main research question is divided into more specific research questions (RQs) for each analysis chapter, as presented in Table 3-1.

Table 3-1: Overview of RQs

Chapter 4	1-1. How is the concept of surveillance defined in the S&S journal? 1-2. Which words are consistently salient across the S&S journal volumes? 1-3. How do the meanings of the consistently salient words shift across the S&S journal volumes?
Chapter 5	2-1. How is the surveillant landscape multimodally represented in concrete examples? 2-2. How does the social dimension – the interaction order – contribute to the textual representation of surveillant landscapes in the Surveillance Blog Corpus compared to the S&S Corpus?
Chapter 6	3-1. How do long-term co-occurrence patterns in the surveillance discourse of <i>The Times</i> develop from 1986–2008? 3-2. Which locally salient patterns are associated with relative frequency peaks in the surveillance discourse of <i>The Times</i> from 1986–2008? 3-3. How does the development of the surveillance discourse relate to newspaper sections in which the nodes occur? 3-4. How is the debate about the Identity Cards Act 2006 reflected in the surveillance discourse of <i>The Times</i> ?

RQs 1-1 to 1-3 are concerned with the representation of surveillance in academic discourse, i.e. in the S&S Corpus, a corpus compiled of the *Surveillance & Society* journal. These questions are addressed in Chapter 4. RQ 1-1 deals with the ways in which surveillance is defined in the journal to gain an initial overview of its meaning. The remaining two RQs focus on how lexical patterns in the corpus contribute to surveillance discourses. RQ 1-2 looks for words that are saliently used across the whole corpus. RQ 1-3 examines shifts in the meaning of these words in relation to main themes identified via a previously published content analysis of the journal (Mehrabov, 2015).

RQs 2-1 and 2-2, addressed in Chapter 5, focus on the coordinate of place (see Section 2.6.1) in surveillance discourse, applying the theoretical framework of surveillant landscapes (Jones, 2017). RQ 2-1 deals with the multimodal representation of surveillant landscapes in concrete examples, that is, in the physical environment of a local shopping centre and online texts about this place. RQ2-2 is concerned with the social dimension (the interaction orders) of surveillant landscapes in the Surveillance Blog Corpus in comparison to the S&S Corpus as a reference corpus.

The final set of RQs, 3-1 to 3-4, examines the representation of surveillance in the Times Digital Archive (TDA) from 1986 to 2008. RQ 3-1 is concerned with tracing co-occurrence patterns in surveillance discourse in the long-term, i.e. the beginning, middle and final years of the corpus (1986, 1997 and 2008). For RQ 3-2, a more bottom-up approach is followed, that first identifies peaks in the relative frequency of the nodes (*surveillance*, *privacy* and *CCTV*) and then examines locally salient co-occurrence patterns around these peaks. RQ 3-3 follows on from the findings for the first two RQs that highlight salient patterns in particular sections of the newspaper. This question therefore examines the diachronic distribution of the nodes across different newspaper sections. Finally, RQ 3-4 focuses on a specific case study of

surveillance discourse in the UK, the debate about the Identity Cards Act 2006, which raised widespread public attention.

The following sections introduce the data and the analytical methodology of this thesis. Section 3.2 describes the corpora and additional datasets underlying the three analysis chapters. All analysis chapters employ the novel methodology of co-occurrence comparisons, which is explained in Section 3.3. Finally, Section 3.4 concludes the chapter.

3.2 Corpora and additional data sources

Meaning-making principle (ii) states that meaning emerges via comparison. In order to follow this principle, it was necessary to gain access to surveillance discourses from different domains. For the analysis in Chapters 4 and 5, specialised corpora have been compiled (of approximately 2.5 – 2.7 million words). The large newspaper corpus for the analysis in Chapter 6 has been processed from a full-text version of the Times Digital Archive (approximately 1.5 billion words) based on particular social events. Each analysis chapter focuses on one of the corpora, respectively.

Figure 3-1 represents the different approaches according to which the corpora of this thesis were compiled. As the curved arrows indicate, the differences in the compilation of the two specialised corpora can be described as a contrast between top-down and bottom-up selection of text (see Section 2.6.2 in relation to temporal segmentation). The S&S Corpus corresponds to the text-external criteria and the selection of the blog posts largely follows a text-internal approach (see Section 2.4.2). The sampling of articles from the Times Digital Archive combines elements from both.

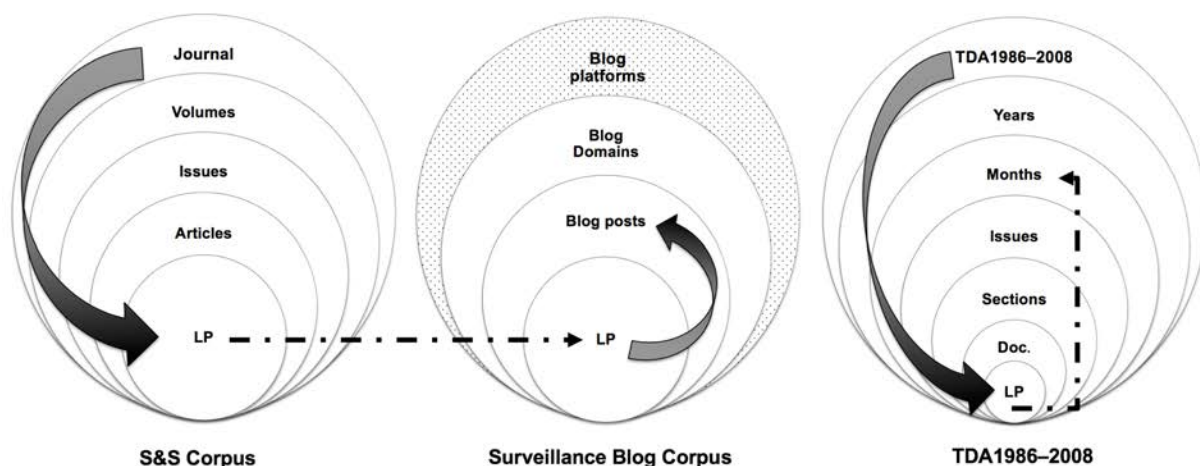


Figure 3-1: Levels of meaning-making in the S&S Corpus and the blog corpus; external vs. lexical criteria for corpus building

The text-external and text-internal criteria can also be understood in terms of what I call the ‘levels of meaning-making’, which are visualised for the three corpora in Figure 3-1. The lowest level that I am concerned within all three corpora is that of ‘lexical patterns’ (“LP” in Figure 3-1). This level takes a different role in the compilation process for the corpora. For the S&S Corpus, the documents are collected based on the outer-most level – texts published in the journal – and the lexical patterns are part of the texts. In Figure 3-1, I use the term ‘articles’ for convenience, as research articles form the majority of the corpus, but some of the texts in the S&S Corpus are editorials or book reviews (see Section 3.2.1). Articles are published as part of the usual quarterly issues that form an annual volume. For mostly practical reasons, the volumes form the main unit of the comparative analysis in Chapter 4 (see Section 3.3.2), but details of the issues and articles are checked as appropriate.

In contrast to the S&S Corpus, the Surveillance Blog Corpus is collected in a bottom-up fashion: starting from the lowest level, the lexical patterns that the retrieved blog posts must contain. In the blog diagram, the outer-most level is dotted, because the posts are restricted to publication on one of the three common blog platforms (Blogspot, Typepad and WordPress), which is explained in Section 3.2.2. This is, however, a purely practical decision in order to

restrict the *BootCaT frontend* tool to collect only blog posts as opposed to other websites. There is no obvious theoretical reason suggesting differences in perceptions of surveillance between the users of these blog platforms. Above the level of the post is the blog domain, i.e. the individual blog that has published a particular post. As the landing page of the blog, the blog domain is an important meaning-making unit, a platform for the content that the blog contains. And although readers may reach a post through a hyperlink from another blog or a search engine result, the blog domain provides the context and a home for the post, potentially pointing to an “About” tab and an archive of all existing posts.

Finally, the meaning-making levels in Figure 3-1 demonstrate that the structure of the TDA1986–2008 is inherently diachronic: the corpus contains 23 years’ worth of data that can be analysed as a whole or as annual subcorpora, which are further subdivided into monthly units. The format of the TDA corpus used in this thesis means that these monthly subcorpora are the smallest units accessible for the corpus linguistic analysis, which I explain in Section 3.2.3. However, the TDA online interface provides an alternative access to the *Times* data, which I use to look up the context of concordance lines in their articles and original layout of the newspaper page and section (see Section 3.2.5). The TDA compilation resembles the top-down selection of the S&S Corpus in that the texts are qualified based on their publication venue, as indicated by the curved arrow in Figure 3-1. At the same time, the corpus selection involves a bottom-up element: for parts of the analysis in Chapter 6, monthly subcorpora are targeted based on peaks in the relative frequency of the node words, indicated by the dashed arrow from the “LP” to the “Months” level in the TDA1986–2008 diagram. This bottom-up process is explained in Section 6.3.

The differences in the discourse domain and composition of the corpora means that they also differ in the degree to which the concept of surveillance is at the core of the debates in the

corpora. The S&S Corpus has the strongest thematic focus on surveillance: the mission of the journal involves the study of surveillance and raising awareness of the concept among the wider public (“Editorial Policies,” 2014; also see Chapter 4). By nature of the lexical compilation process it is ensured that all texts in the Surveillance Blog Corpus contain some lexical features (bigrams) that relate to surveillance discourse. As I explain in Section 3.2.3, these “surveillance bigrams” are derived from the S&S Corpus. Hence, a dashed arrow leads from the LP level of the S&S Corpus to the LP level of the Surveillance Blog Corpus. Nevertheless, a corpus of lexically gathered blog posts is likely to be much more diverse than a corpus of texts that all derive from the same academic journal. Lastly, as a large, general news corpus, the TDA1986–2008 contains discussions of a wide range of current events and topics of news value in the covered period. A central assumption of this thesis is that surveillance is an important social issue (see Section 2.2). As newspapers discuss events and issues that involve news values (see Section 2.5.3), *The Times’s* coverage from the 1980s to 2000s can be expected to refer to the social issue of surveillance. The following sections describe the collection criteria and main attributes of the three corpora.

3.2.1 Academic journal articles: The S&S Corpus

Chapter 4 follows a similar approach to that of Teubert (2007a) by examining a corpus that is self-contained. The corpus contains all volumes of the *Surveillance & Society* (S&S) online journal from its launch in 2002 until 2015, representing an authentic dataset that was, at the time of compilation in September 2015, complete. By compiling a corpus from the S&S journal, I can make use of the general advantages of analysing the discourse from externally-defined corpora. First, the corpus compilation is not biased through the use of search terms often employed in the compilation of specialised newspaper corpora (see Section 2.4.2.2). Second,

an entire discourse can be studied; that of research articles in this surveillance-focused journal. In the following sections, I explain the rationale for the text-external compilation criteria (3.2.1.1), the retrieval and cleaning of the corpus files (3.2.1.2), and the specifications of the finalised corpus S&S Corpus (3.2.1.3).

3.2.1.1 Text-external criteria

The S&S Corpus has been compiled according to the text-external criterion of publication venue. This measure interacts with the approach of compiling corpora based on the recommendations of subject experts (see Section 2.4.2.1). I consider the publication in a specialised open-access journal as evidence for expert endorsement. The S&S editorial team positions the journal as “the premier journal of surveillance studies” (“About the Journal,” 2019, online). The website further sets out the focus and scope of the journal with the following aims (“About the Journal,” 2019, online):

- publish innovative and transdisciplinary work on surveillance;
- encourage understanding of approaches to surveillance in different academic disciplines;
- promote understanding of surveillance in wider society;
- encourage policy and political debate about surveillance.

The high profile of the S&S journal in the field of surveillance studies means that the journal is a suitable dataset for analysing specialist surveillance discourses. The meaning-making levels of the S&S Corpus in Figure 3-1 directly relate to the publication process of the journal. A detailed overview of the composition of the corpus is given in Table 3-2.

The raw corpus is article-based with each file containing a single document (e.g. an individual editorial, research article or book review), but arranged in volume directories. Most of the analysis of the S&S Corpus works with the 13 volume subcorpora (see Section 3.3.2). The file names are unique identifiers, structured as “Year-Volume-Issue-Article No.”. For example, the file “2015-13-2-17” contains the 17th document (a research article) in Volume 13 (Issue 2), published in 2015. Appendix A lists all included articles with their identifiers, authors and titles. Examples that I cite in the text are additionally listed in a bibliography for the S&S Corpus, separately from the bibliographies for examples from the other two corpora.

Table 3-2: List of S&S issues

Volume	Issue	Year	Issue title	Files	Tokens ¹
1	1	2002	<i>Launch issue</i>	34	190,471
1	2	2003	Work		
1	3	2003	Foucault and Panopticism Revisited		
1	4	2003	Surveillance and Mobilities		
2	1	2004	<i>Open Issue</i>	37	246,572
2	2, 3	2004	The Politics of CCTV in Europe and Beyond		
2	4	2004	People Watching People		
3	1	2005	<i>Open Issue</i>	17	104,598
3	2, 3	2005	Doing Surveillance Studies		
4	1, 2	2006	<i>Open/ Conflict</i>	23	131,071
4	3	2007	Surveillance and Criminal Justice: Part 1		
4	4	2007	Surveillance and Criminal Justice: Part 2		
5	1	2008	<i>Open Issue</i>	33	133,558
5	2	2008	Smart Borders and Mobilities: Spaces, Zones, Enclosures		
5	3	2008	Surveillance and Inequality		
6	1	2009	<i>Relaunch Issue: Revisiting Video Surveillance</i>	72	208,132
6	2	2009	Health, Medicine and Surveillance		
6	3	2009	Surveillance and Resistance		
6	4	2009	Gender, Sexuality and Surveillance		
7	1	2009	<i>Open Issue</i>	26	140,544
7	2	2010	Surveillance, Performance and New Media Art ²		
7	3, 4	2010	Surveillance, Children and Childhood		
8	1	2010	<i>Open Issue</i>	62	279,325
8	2	2010	Surveillance and Empowerment		
8	3	2011	Marketing, Consumption and Surveillance		
8	4	2011	<i>Open Issue</i>		
9	1, 2	2011	A Global Surveillance Society?	38	224,013
9	3	2012	Urban Surveillance		
9	4	2012	Cyber-Surveillance in Everyday Life		
10	1	2012	Surveillance in Latin America	41	181,915
10	2	2012	<i>Open Issue</i>		
10	3, 4	2012	<i>Open Issue</i>		
11	1, 2	2013	Surveillance Futures	43	259,357
11	3	2013	Surveillance Texts & Textualism: Truth-telling and Trustmaking in an Uncertain World		
11	4	2014	Surveillance and Sport		
12	1	2014	<i>Open Issue</i>	56	293,048
12	2	2014	Big Data Surveillance		
12	3	2014	Surveillance, Gaming and Play		
12	4	2014	<i>Open Issue</i>		
13	1	2015	Doing Surveillance Studies (2)	30	153,385
13	2	2015	Surveillance and Security Intelligence After Snowden (Part 1)		
Total				512	2,545,989

¹ These are the tokens given in the *WordSmith Tools* WordList “statistics” tab for the tokens “used in wordlist”, i.e. here the counts exclude numbers. I adjusted the setting for symbols to be allowed in tokens so that they can contain apostrophes and periods. The periods ensure that abbreviations like *e.g.* and *i.e.* are not split into individual letters (but *WordSmith Tools* will remove the final period, so display only *e.g.* without in the lists).

² Most of the documents from this issue have been excluded because they are ‘artistic presentations’, leading to a relatively small subcorpus size of Volume 7.

Most S&S issues are special issues with a particular thematic focus. Usually a special issue consists of an editorial and a number of articles related to the issue's topic and several book reviews. There are links across issues in individual cases, for instance when a thematic issue is split into two issues such as Issues 4(3) and 4(4), as shown in Table 3-2. Out of the 41 issues in the corpus, 11 "open" issues contain various topics (including the "launch issue"). Some issues are part "special" and part "open", such as Issue 4(1/2) ("Open/Conflict"; see Table 3-2). Issue 6(1), titled the "Relaunch" issue, marks the move to a new website. The list of issue titles indicates that the journal covers a large range of themes related to surveillance, including theory, like 1(3) "Foucault and Panopticism Revisited", methodology, e.g. 3(2/3) "Doing Surveillance Studies", and particular applications, such as 11(4) "Surveillance and Sport". Chapter 4 discusses some of these differences as part of the co-occurrence comparison across volumes and in relation to Mehrabov's (2015) content analysis findings.

3.2.1.2 Text retrieval and corpus cleaning

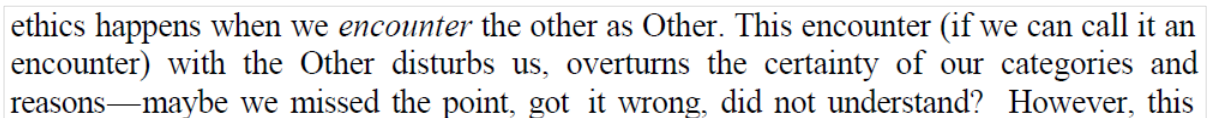
This section outlines procedure followed to retrieve and clean texts for the S&S Corpus. All information on the S&S journal structure and its articles listed here is based on the journal website (<http://www.surveillance-and-society.org>). Since S&S is an open-access journal, article files in PDF format are freely accessible from its website, "[I]icensed to the Surveillance Studies Network under a Creative Commons Attribution Non-Commercial No Derivatives license".³

The PDF files were downloaded and converted to plain text format for the corpus linguistic analysis. For the conversion I used the software *Solid Converter Mac (Trial Version)*

³ The license is available from <https://creativecommons.org/licenses/by-nc-nd/3.0/>

(Solid Documents Limited, 2015), which automatically removed most headers and footers. Manual cleaning was necessary to remove any remaining headers and footers, horizontal lines and column separators that would be read as special symbols in concordance software. During this data cleaning stage, I also manually removed footnote indicators (asterisks and numbers) and text. As footnotes interrupt the body of text, they would skew the results of corpus linguistic analyses such as collocation.

A small number of duplicate files and missing files were encountered. These issues are documented in Appendix B. As the main focus of this project is on the corpus analysis rather than corpus compilation and pre-processing, a balance had to be struck between the time and effort dedicated to these stages. I aimed to remove duplicates and clean metadata like footnotes, captions and bibliographies. These complications create a relatively laborious compilation process even with an “orderly” data source like this academic journal which at least has a gold standard. For example, a given article is clearly associated with a particular volume/issue number and if it appears twice this is wrong. As I show in Section 3.2.2, this situation is more complex with a “messier” dataset like blog posts. When checking the S&S Corpus data, I also noticed some inconsistencies in spacing and hyphenation, probably introduced during the conversion from PDF to plain text. I identified 2,600 cases where a dash (—) was used without spaces (with the regular expression `\b—\b`) like *reasons—maybe* shown in Figure 3-2. In some programmes this leads to falsely hyphenated words (e.g. *reasons-maybe*), as the tokenisers for *WordSmith Tools* (Scott, 2016) and the *quanteda R* package (Benoit & Nulty, 2016) used for the co-occurrence comparison for the analysis in Chapter 4 recognise the dash as a hyphen. Using the text editor *TextWrangler* 5.5.1 (Bare Bones Software, 2016) I replaced the word boundaries in the regular expression with spaces. In addition, I replaced occurrences of double hyphens (--) without spaces with single hyphens surrounded by spaces.

A screenshot of a text document showing a line of text: "ethics happens when we encounter the other as Other. This encounter (if we can call it an encounter) with the Other disturbs us, overturns the certainty of our categories and reasons—maybe we missed the point, got it wrong, did not understand? However, this". The text is enclosed in a light blue rectangular box. The hyphenation error is visible at the end of the line, where the text is cut off by the box boundary.

ethics happens when we *encounter* the other as Other. This encounter (if we can call it an encounter) with the Other disturbs us, overturns the certainty of our categories and reasons—maybe we missed the point, got it wrong, did not understand? However, this

Figure 3-2: Screenshot illustrating the use of a dash leading to falsely hyphenated words⁴

Some other errors were only recorded at this stage, but not resolved. Another special type of long dash character was, for example, found in article 2012-10-1-05. *WordSmith Tools* does not recognise this dash, but pictures it as a square box, leading to some unique, wrong tokens (e.g. “institutions□□is” instead of “institutions—is” or “institutions – is”). This appears to be a small number of errors and is related to the challenge of cleaning the encoding of converted files. Additionally, hyphenation errors are introduced “when words are split during typesetting, leaving a hyphenated word at the end of a line which continues in the next line. The *de-hyphenation* of such forms is not trivial” (Schmid, 2008, p. 531, emphasis in original). A search for end-of-line hyphens showed that their impact is likely to be small, with 179 hyphens and 47 dashes.

The S&S journal includes different types of academic writing. For the purposes of the present study, the more traditional genres such as editorials, research articles and book reviews are of main interest. According to the editorial scope section on the S&S website, the journal “encourages submissions that could not be published in conventional paper journals such as html, photographic, video and new media work” (“Editorial Policies,” 2014). These more creative genres found in some issues of S&S – two poems, three interviews and 11 artistic presentations – have been excluded from the S&S Corpus. The rationale here was (i) to collect only textual data as accounting for multimodal performance submissions was beyond the scope

⁴ Example taken from article 2003-01-2-06 (Introna, 2003, p. 211)

of this study; (ii) to keep the types of data within the corpus relatively consistent, and (iii) to represent academic discourses of surveillance. Most of the artistic presentations are actually part of Issue 7(2), the special issue on performance and new media art, which is therefore an outlier compared to the rest of the corpus.

3.2.1.3 Specifications of the finalised S&S Corpus

The breakdown of the document categories in the S&S Corpus is shown in Table 3-3. More than half of the files are articles. As the RQs for Chapter 4 are concerned with the discursive representation of surveillance in the corpus overall, no further distinctions are made between the different categories.

Table 3-3: Distribution of categories in the S&S Corpus

Category	Number of files
Article	274
Book review	135
Editorial	28
Opinion/view	25
Debate	20
Review article	15
Case study	10
Research note	4
Presentation	1
Total	512

The plain text files were the basis of the analysis for Chapter 4. Section 3.3.2 outlines the specific analysis procedure of the cross-volume comparison in Chapter 4. In Chapter 5 I return to the S&S Corpus to employ it as a reference corpus for the Surveillance Blog Corpus that is the focus of the next section. The analysis framework that I developed for comparing these two

corpora in Chapter 5 required another format of the data, using semantically tagged corpora. In Section 3.2.4, I outline how the corpora were tagged.

3.2.2 The Surveillance Blog Corpus

For the analysis of public surveillance discourse in blogs, another corpus was compiled based on text-internal criteria. The description of the compilation process for the Surveillance Blog Corpus starts with the rationale for the text-internal criteria and their implementation (3.2.2.1), before moving on to the retrieval of the corpus texts from the internet (3.2.2.2) and cleaning procedures undertaken after the retrieval (3.2.2.3). Section 3.2.2.4 presents the specifications of the finalised corpus.

3.2.2.1 Text-internal criteria

In Chapter 5, the lexical focus of the thesis is taken further by analysing a specialised corpus that is built according to lexical criteria, that is, a list of ‘seed’ words. These seeds are taken from the externally defined S&S Corpus introduced in 3.2.1. Chapter 5 argues that a specialised corpus based on situational or other external criteria can be usefully employed as a ‘seed corpus’ for lexical patterns that are then used as the defining features for building corpora from a more “flexible” background. So, the external criteria of the S&S Corpus – i.e. peer-reviewed academic writing about surveillance, selected for publication in this journal – restrict the contents of that corpus in terms of style and focus. For documents from different platforms on the internet an external categorisation is not straightforward, because they do not readily show the same “belonging” to a particular background as the documents published in the S&S journal. However, I can simplify this procedure by making use of the specialist surveillance corpus to define text-internal criteria for another specialised corpus. Here, the internal criteria

to be employed are lexical. Chapter 2 has argued that existing approaches to search term selection have focused more on practical than theoretical questions. The present study seeks to combine a theoretical perspective with an operational approach to search term selection. As a result, the methodology for Chapter 5 puts forward an approach to the selection of search/seed terms for a corpus not based on a rough trial corpus, but a carefully compiled specialised ‘seed corpus’.

3.2.2.2 Text retrieval

The compilation of the Surveillance Blog Corpus was carried out using *BootCaT frontend* Version 0.71⁵ (Baroni & Bernardini, 2004; Zanchetta, Baroni, & Bernardini, 2011; henceforth *BootCat*), a user-friendly tool for compiling specialised web corpora (see Gatto, 2014). Like other web crawling approaches, *BootCaT* relies on a set of ‘seed’ terms in the retrieval of URLs. The selection of seed terms for corpus compilation is not straightforward; various factors have to be balanced, including relevance, objectivity and restrictions from the database (see e.g. Gabrielatos, 2007). Baroni and Bernardini (2004) suggest a recursive procedure; first retrieving documents for a pilot corpus with a given set of seed terms, then retrieving more documents based on the keywords of this pilot corpus. The present study introduces a variation of this procedure to employ a fully compiled seed corpus instead of a rough pilot corpus of websites: it uses bigrams formed by salient words in the S&S Corpus as seed words. Employing the S&S Corpus as a seed corpus for the blog corpus has the advantage that the seed terms have a meaningful source. Thus, seed terms cannot be influenced by any potentially irrelevant documents retrieved for a pilot corpus.

⁵ Retrieved from <http://bootcat.dipintra.it/?section=home> (accessed October 2016)

The *BootCaT* URL retrieval was run once for each of the three blog platforms, because the present study only requires a moderately sized corpus that is comparable in size to the S&S Corpus. The retrieval of the corpus data consisted of four main steps, (i) selecting seed terms; (ii) generating random combinations of seed terms, called ‘tuples’ in *BootCat*; (iii) collecting URLs and (iv) retrieving full texts, as I explain in the following subsections.

(i) Selecting seed terms

The aim in selecting seed terms from the S&S Corpus was to identify a list of terms that describe surveillance discourse, but are not specific to academic writing. Bigrams were used as seed terms, as done by Elgesem and Salway (2015), because these tend to be more specific than single words and can help alleviate ambiguity of polysemous words. For this purpose, I generated a list of those bigrams in the S&S Corpus that contain S&S key keywords (see Sections 3.3.2 and 4.3) using the *WordSmith Tools 7* (Scott, 2016) function “words to make clusters from a text-file”. These key keywords are a useful starting point because they are by definition used frequently throughout the entire S&S Corpus and tend to represent important concepts. The resulting list contained 2,643 bigrams, which would be too many to use as seed terms. The more terms are used, the larger and potentially the more heterogeneous the corpus. Baroni and Bernardini (2004) report that 5–15 seed terms can be sufficient. In this study I used 60 because the aim was to narrow down the list of candidate bigrams in a principled way rather than introducing my own bias in choosing the terms. A set of quantitative and qualitative criteria was established for bigrams to qualify as useful seed terms:

- frequency ≥ 50 – excluded 2,390 items;
- dispersion ≥ 10 texts – excluded three additional items;

- content words only – excluded 174 additional items (e.g. *surveillance and, forms of, surveillance in*);
- not too generic – excluded eight additional items (e.g. *everyday life, social life, social groups*); and,
- not too specific to academic fields and academic writing (e.g. *surveillance studies, social science, disciplinary powers*) – excluded eight additional items.

Table 3-4: Seed bigrams for the blog corpus

Rank	Bigram	Freq.	No. of Texts	Rank	Bigram	Freq.	No. of Texts
1	surveillance technologies	785	201	31	cctv operators	93	30
2	surveillance practices	526	170	32	enforcement agencies	92	41
3	data protection	472	80	33	social network	91	40
4	surveillance systems	445	153	34	social order	85	44
5	social control	411	140	35	privacy advocates	80	38
6	social media	358	55	36	social justice	78	33
7	surveillance society	313	125	37	surveillance measures	77	43
8	data collection	294	103	38	cctv footage	77	26
9	surveillance cameras	290	84	39	social security	76	31
10	cctv cameras	280	77	40	monitoring systems	74	20
11	cctv systems	252	64	41	security staff	73	19
12	cctv surveillance	208	44	42	security agencies	70	33
13	social networking	176	52	43	surveillance mechanisms	68	39
14	surveillance technology	172	91	44	contemporary society	65	39
15	surveillance system	154	68	45	surveillance apparatus	65	29
16	contemporary surveillance	153	79	46	digital technologies	63	30
17	social sorting	150	73	47	disciplinary society	63	24
18	camera surveillance	145	33	48	surveillance activities	62	37
19	cctv system	142	46	49	surveillance data	62	28
20	social relations	129	68	50	cctv camera	61	27
21	data mining	121	49	51	data processing	61	22
22	surveillance techniques	118	64	52	identification system	61	13
23	social networks	113	40	53	social exclusion	60	22
24	social surveillance	111	20	54	surveillance practice	57	27
25	surveillance camera	105	46	55	data gathering	53	29
26	privacy rights	100	51	56	privacy concerns	53	25
27	security personnel	98	31	57	surveillance devices	53	24
28	security guards	98	30	58	internet users	52	22
29	security measures	97	48	59	surveillance tools	50	28
30	privacy protection	96	40	60	digital surveillance	50	26

Table 3-4 shows the resulting 60 seed terms, which meet the above criteria, with their frequencies and dispersions in the S&S Corpus. The rationale behind this selection of seed

terms was to create a specialised set of seed terms in order to prevent the collection of irrelevant texts.

(ii) Automatic generation of three-word tuples to crawl relevant texts

BootCaT automatically combines the seed terms into random sequences called ‘tuples’ used to query the web. For this study, the length of tuples was set at three, i.e. each tuple contained three bigrams from Table 3-4, which is the recommended length for a specialised corpus (Zanchetta & Gaspari, 2018). Based on trial and error, I chose to create 200 tuples in order to compile a corpus of similar size to the S&S Corpus. The same list of tuples was used for querying all three blog platforms, starting with the following tuple: “surveillance data” “surveillance technology” “cctv camera”. A blog post in the corpus can still contain any of the phrases that do not meet these criteria; the tuples are just used as the search query for identifying texts.

(iii) Automatic collection of matching URLs from Blogspot, Typepad and WordPress

This step was carried out separately for each of the three blog platforms using *BootCaT*, but the same list of tuples from Step (ii) was used in each case. In the first collection of URLs, the website domain was restricted to Blogspot (.blogspot.com), in the second collection to Typepad (.typepad.com) and, finally, to WordPress (.wordpress.com).⁶ *BootCaT* does not search the internet directly but relies on a generic search engine to identify relevant URLs. The corpora for this study were compiled with *BootCaT* version 0.71 running a Bing search in October 2016, which required a key from the Windows Azure Marketplace. This procedure has changed with later versions of *BootCaT*.

⁶ Alternative UK domains (ending in “.co.uk”) were accessed in a trial study, but not included in the final data collection, because some blog posts retrieved from the UK domain in a trial run overlapped with those found in the “.com” collection, particularly in the case of Blogspot addresses.

(iv) *Retrieving full texts for each of the three blog platforms*

In this step, *BootCaT* accessed the identified URLs for the full text retrieval, which was mostly successful. Even for the blog platform with the biggest difference between identified URLs and successful text retrieval in this study (Blogspot), a success rate of approximately 88% was achieved (438 URLs retrieved out of 498 identified URLs). The output from *BootCaT* consists of one long text file for each collection procedure, i.e. in this study one for each blog platform. In each file, the text chunks retrieved from distinct URLs are marked with the phrase *CURRENT URL* in the text files⁷, as illustrated in Example (1), which shows the beginning of one blog post.

- (1) *CURRENT URL* <http://150patrick44.blogspot.com>
Patrick's J150 ePortfolio
Saturday, December 14, 2013
With modern surveillance techniques and technologies, it is possible to physically go missing but impossible to disappear digitally
We are being digitally tracked and monitored today more than ever before.
(blogspot_outfile000; Patrick, 2013)

3.2.2.3 Corpus cleaning

According to the *BootCaT* website, the software has “full UTF-8 support” (“Bootcat:release_notes:0.71 [Docs],” n.d.). Although the corpus contains some noise from internet texts (including special symbols for other languages and layout such as bullet points), cleaning this is not the focus of the present study. Unlike the S&S Corpus, which was relatively rigorously cleaned (see Section 3.2.1.2), the Surveillance Blog Corpus is not cleaned in as much detail because of its inherent heterogeneity caused by the large variety of source blogs and the

⁷ The marker *CURRENT URL* is therefore part of the corpus data. As there is only one per file, the impact on the analysis is limited.

consequent lack of a “gold standard” for comparison. So, for the blog corpus, no clear standard exists in terms of what elements should be “cleaned”, with different blogs containing different formats of the boilerplate information etc.

For an easier overview of individual entries, the files were divided at the *CURRENT URL* marker shown in Example (1). This division was completed using the `csplit()` function on the Mac terminal. Accordingly, each individual corpus file contains only the URL, its marker and the text that *BootCaT* has received from the respective URL.

Table 3-5 summarises the stages of the cleaning procedure for the blog files, listing the number of files removed and remaining at each stage. The removal stages in Table 3-5 add up to a total of 493 files that were removed, leaving 346 remaining files in the final corpus. The following paragraphs explain the stages of this cleaning procedure.

Table 3-5: Overview of cleaning procedure for the Surveillance Blog Corpus

Cleaning stage	No. of files removed	Files remaining
1. Initial retrieved files	/	839
2. Filtering for <i>surveillance</i>	- 444	395
3. Random sample check of ~10% (40 files)	/	395
4. Remove files exclusively containing lists of links/headings/(book)references and/or appear to be spam (based on visual inspection)	- 21	374
5. Remove duplicate posts	- 28	346

Stage 1 refers to the retrieval of the initial 839 files. No files were removed at this stage. Stage 2 involved filtering for the word *surveillance*, so all texts that did not contain the word *surveillance* (as an individual token or part of a hyphenated compound⁸) were excluded. This stage was motivated by the observation from manual checks of selected files that not all files

⁸ A simple `grep` search was used for the sequence *surveillance*: `grep -ilr "surveillance"` (where the “-i” flag means case-insensitive; “-l” shows only matching files; “-r” searches in a directory). This search also matched two (out of 395 files) based on hyphenated compounds: *self-surveillance* (typepad_outfile_014) and *camera-surveillance* (typepad_outfile_058).

seemed to contain full tuples used in the retrieval procedure. Spot checks of the live websites showed that in these cases the search terms tended to appear somewhere on the website, but outside the actual post (e.g. in the tags, the side columns or in another post). Therefore, the tuple terms were not retrieved by *BootCaT*. This reduced the number of blog files by slightly more than half, from 839 to 395.

In Stage 3, a random sample of approximately 10% per subcorpus of the remaining 395 “filtered” files was checked manually in comparison with the online presentation. So, for 40 files, I visited the original URL and compared the online text to the text file. The rates of files containing multiple posts (ten files; 25% of the sample) and files containing shortened posts (seven files; 17.5% of the sample) are relatively low. Files seem to contain multiple posts when the URL captured by *BootCaT* points to the blog’s landing page or a link that shows a collection of posts (for a particular month or thematic tag). Conversely, the sample files containing shortened posts tend to originate from blog posts that contain special formatting and/or are extremely long, potentially causing *BootCaT* not to retrieve the entire post. Out of the 40 files, 28 (70%) contained some form of boilerplate. This was usually limited to phrases like “posted by” or the date. Overall, the findings of these formal criteria were judged to be satisfying for the corpus. However, a more concerning observation of the random sample check was that several posts appeared to contain more noise than meaningful contribution to the discourse. An example is given in Figure 3-3. This post contains seemingly unrelated listings of phrases and links.



Figure 3-3: Screenshot of apparent noise in the ‘t0mear’ blog post (blogspot_outfile369)⁹

It is difficult to judge for every case whether a text is “genuine” or “spam”, without detailed contextual research (and even then it is problematic to make such a value judgement). Based on the potential “spam” examples found in the random sample, I decided to carry out a systematic visual inspection of text files. So, in stage 4, I visually inspected all remaining text files and removed any files that appeared to only contain lists of links, headings, bibliographical references and/or raise the impression to be “spam” (as in Figure 3-3). I removed 21 files at this stage (see Table 3-5).

Finally, in stage 5, I removed 28 files that appeared to be duplicates of another file (see Table 3-5). *BootCaT* had already removed any duplicates that had the same URL. It is not straight forward to identify true duplicates of entire files. So, I considered files to be duplicates if the first four lines were identical.

⁹ I do not provide the URL for this instance, because this appears to be an uncontroversial instance of “spam”.

3.2.2.4 Specifications of the finalised Surveillance Blog Corpus

The specifications of the final Surveillance Blog Corpus are given in Table 3-6. With approximately 2.7 million words it is slightly larger than the S&S Corpus. The Blogspot component is by far the largest, in terms of the files included and the number of tokens. The 346 files in the final corpus, listed in Appendix F, originate from 255 different blogs (unique domains). Accordingly, a small number of blogs are represented more than once in the corpus. As Table 3-7 shows, the distribution seems almost Zipfian: one blog domain yields 18 files and then the number of files rapidly decreases, with 219 blogs corresponding to only one file. The top blog domain, <http://amberhawk.typepad.com>, points to the legal blog *Hawktalk* (Amberhawk, n.d.) written by the law training business Amberhawk Associates (located in West Yorkshire, UK).

Table 3-6: Breakdown of corpus size

(Sub)corpus	No. of files in final corpus	Final corpus size in tokens
Blogspot	201	2,108,530
Typepad	27	115,308
WordPress	118	513,397
TOTAL	346	2,737,235

Table 3-7: Breakdown of files per blog domain

Blog domain	Number of files
http://amberhawk.typepad.com	18
https://paulbernal.wordpress.com	10
http://consumercal.blogspot.com	8
https://privacynewshighlights.wordpress.com	7
http://worldunderwatch.blogspot.com	5
https://ubisurv.wordpress.com	5
http://antifascist-calling.blogspot.com	4
http://nexusilluminati.blogspot.com	4
https://alethonews.wordpress.com	4
https://medtextfree.wordpress.com	4
https://thinkingshift.wordpress.com	4
[4 different URLs]	3
[21 different URLs]	2
[219 different URLs]	1

Apart from the blog domains, many URLs contain the year, month and (sometimes) the date of posting. For example, the URL in (2) indicates that the time of posting is October 2010. This information is not available from all URLs, for example for those that lead to a collection of posts tagged with a certain theme. Nevertheless, the URLs can provide a useful approximation of the temporal distribution of the posts in the corpus (see Elgesem & Salway, 2015, who use this approach).

(2) <http://amberhawk.typepad.com/amberhawk/2010/10/project-champion-report-misses-the-target-when-discussing-cctv-surveillance-privacy-and-data-protection.html>

It was possible to extract the year from approximately 52% of the URLs (181 out of 346). Table 3-8 lists the number of URLs associated with each year from 2005, the earliest match, to 2016, when the Surveillance Blog Corpus was compiled. Among these, recent years feature more URLs: the entries dated for the years 2013–16 alone correspond to approximately 56% of the URLs featuring years. The scarcity of URLs until 2008 in particular is not surprising, as blogging technology was still relatively recent at the time (see Section 2.5.2). It is also possible that some blogs from the earlier years have since been deleted and could therefore not be retrieved for the compilation.

Table 3-8: Number of URLs associated with a particular year (excluding the 155 URLs without year indication)

2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	Total
3	/	8	9	19	21	13	16	25	26	26	25	181

The Surveillance Blog Corpus was saved in plain text files in the first instance, with one file per URL retrieval (i.e. mostly corresponding to an individual post). For the analysis in Chapter 5, a semantically tagged version of this corpus was created (see Section 3.2.4).

3.2.3 The Times Digital Archive, 1986–2008

Chapter 6 is based on the full text of all issues in *The Times* (of London) from 1986 to 2008, using a local copy of the Times Digital Archive, which has been digitised from newspaper scans and can be accessed online via a subscription service at <https://www.gale.com/uk/c/the-times-digital-archive>. In the following three subsections, I explain how this data was accessed and processed, and I outline the specifications of the final corpus.

3.2.3.1 The CCR local copy of the complete TDA1785–2008

The local copy of the Times Digital Archive (TDA) held at the Centre for Corpus Research, University of Birmingham has been processed to facilitate corpus linguistic analysis by (i) stripping the files of all XML tags surrounding individual words and (ii) dividing each file containing a daily issue into smaller files by article.¹⁰ The XML tags at the beginning of files remain, containing metadata such as the date, article category and title. Paragraphs are marked up in XML tags throughout the tags. Example (3) shows the header and beginning of an article. The XML tags provide various types of information: the encoding of the file, ID numbers of the article and issue as well as the publication date (1st February 2008) and the article category (“News”). Only the article title (surrounded by <title></title> tags) and the main text (within paragraph tags, <p></p>) forms the basis of the linguistic analysis in Chapter 6. The TDA1986–2008 corpus created for this thesis contains the entire text of the newspaper (beyond the “News”, including text such as adverts, TV programmes, sports commentary etc.). However, Section 6.4 examines frequency patterns within different categories using the online interface of the TDA (introduced in Section 3.2.5).

¹⁰ This work was carried out by Dr Anthony Hennessey.

(3) <?xml version="1.0" encoding="UTF-8"?>
 <article issue_id="0FFO-2008-0201" issue_date="2008-02-01"
 article_id="0FFO-2008-0201-0008-003"
 category="News">
 <title>Children find themselves in the line of fire as gun deaths of under-16s
 rise</title>
 <p>Richard Ford Home Correspondent One in twelve victims of gun-related
 murder last year was a child, official figures published yesterday say. Five of the 59
 people shot dead in 2006-07 were aged under 16 com-</p>

3.2.3.2 Processing the data to create the TDA1986–2008 corpus

As this thesis is concerned with surveillance discourses of recent decades, I have chosen to focus on the final period from 1986-2008 of the local TDA files. This period is the most recently digitised section of the TDA. The company that created the TDA has confirmed to me that the data from 1986 onwards is considered to contain a higher OCR accuracy than the 1785–1985 material (S. Cripps, GALE Cengage Customer Care and Technical Support Executive, personal communication, November 28, 2018).

So, for this thesis, I created a subcorpus of the complete TDA covering the years 1986–2008, in a suitable format for my analysis of monthly files. This subcorpus covers all categories of the newspaper, as this project is concerned with the representation of surveillance within the entire newspaper. The OCR scanning of adverts, tables and images may introduce extra noise. However, I did not want to introduce any arbitrary sampling decisions by filtering for categories, as it is not clear how precise the categorisation is. Chapter 6 shows that sections beyond the “News” category actually contribute to the meaning of *surveillance*.

In order to run the co-occurrence comparisons, the XML files had to be converted to a format that is readable by *CorporaCoCo* (dev. version 1.1), the package used for comparing co-occurrences (see Section 3.3.1). The required format is the so-called ‘corp_text object’ that

holds information on the full text, with the tokens in their original format (i.e. with the right capitalisation) and types saved as separate layers. With this format, *CorporaCoCo* can generate co-occurrences counts of types, while linking back to the actual tokens in the original text so that concordances can be computed. During the import process, the default *CorporaCoCo* (dev. version 1.1) tokenisation was used, which implements the Unicode standard using the *stringi* package (Gagolewski, Tartanus, IBM, Unicode Inc., & other contributors, 2017). Unlike the tokenisations used for the analyses of Chapters 4 and 5 (cf. Section 3.3.1), respectively, the Unicode standard considers hyphens as word boundaries. Given the large size of the TDA1986–2008 (see Section 3.2.3.3), the main advantage of running the concordances directly in *R* (R Core Team, 2016) on a powerful server (rather than in a standard concordancer like *WordSmith Tools*) is the speed. The conversion into `corp_text` objects required the following steps using *R*:

- (i) convert individual article files into files containing all text of a daily issue (saved as “text_collapsed[date].rds”);
- (ii) create `corp_text` objects for the daily issues (saved as “[date]_corp_text.rds”);
- (iii) combine daily `corp_text` files to monthly `corp_text` files (saved as “monthly_object[year-month].rds”).

Steps (i) and (ii) were carried out together and the creation of monthly files in step (iii) was carried out after (i) and (ii) had been completed. Due to the large number of daily files for the 23-years’ worth of data and the long time required for each loop of data reading, the entire conversion code for steps (i) to (iii) was run on the University of Birmingham’s BlueBEAR

High Performance Computing system based on code that I supplied to the Research Software Group.¹¹

With the completion of step (iii) the whole dataset has been converted into 276 monthly corp_text object files. This is a more manageable number of files (even though each of them is still rather large) that I was able to process further using the CCR server. The next step was to create frequency counts of each month, for all tokens and particularly for *surveillance* in order to determine peaks in the usage of this word.

3.2.3.3 Specifications of the finalised TDA1986–2008 corpus

Table 3-9 illustrates that the entire TDA1986–2008 can be considered a “mega” corpus with a size of over 1.5 billion words. By comparison, the Corpus of Contemporary American English (Davies, 2008), 1990–present contains >560 million words and is therefore smaller – although of course even larger corpora now exist based on web data. Within the context of the present study, the TDA1986–2008 is a large corpus indeed, as even the smallest monthly subcorpus is bigger than either the S&S Corpus or the Surveillance Blog Corpus.

Table 3-9: Token counts in the TDA1986–2008

Total token count	1,515,167,653
Average monthly token count	5,489,738
Minimum monthly size	3,396,244 (in 1986_02)
Maximum monthly size	7,286,901 (in 2004_01)

¹¹ Support was provided by Dr Simon Branford, of the Research Software Group, Advanced Research Computing, University of Birmingham. The RSG provide research software services to the University’s research community. See <http://www.birmingham.ac.uk/bear-software> for more details.

3.2.4 Semantically tagged versions of corpora for Chapter 5

Chapter 5 aims to go beyond the lexical results of individual nodes and collocates in order to examine wider semantic co-occurrence patterns. To achieve that, I created semantically tagged versions of the two corpora compared in that chapter: the Surveillance Blog Corpus (see Section 3.2.2), which the analysis focuses on, and the S&S Corpus (see Section 3.2.1), which acts as a reference corpus in Chapter 5.

The *Constituent Likelihood Automatic Word-tagging System* (Garside, 1987; *CLAWS*) and *USAS* tagger (Archer et al., 2002; see Section 2.4.6), on which the *Wmatrix* online tool is based, were used for (i) POS-tagging and (ii) the semantic annotation. An offline version of these taggers was used, because the corpora were larger than the recommended size for *Wmatrix* and it was useful to keep individual corpus files separate (*Wmatrix* requires one input file).¹² In order to ensure that the taggers would run smoothly, it was necessary to make additional adjustments to the texts. According to the input format instructions on the *CLAWS* website (“*CLAWS* Input / Output Format Guidelines,” n.d.), the texts were processed with a Perl script to replace special characters with entities that the tagger would understand. Each corpus file was first tagged with the *CLAWS* tagger and secondly with the *USAS* tagger. The tagged files were saved as ‘corp_text’ objects for the co-occurrence comparison with the development version of *CorporaCoCo* 1.1.¹³ As I show in Section 3.3.1, the corp_text objects encapsulate the text and the tokenisation of a corpus. For the semantically tagged corpora this feature was of particular value, because it makes it possible to directly compare the co-occurrences of semantic tags, while also retaining the original text for analysing the results in the context of concordance lines. *USAS* sometimes finds matches for several tags that are then listed in the

¹² Both taggers (unversioned) were kindly provided by Paul Rayson on 21 December 2017.

¹³ The offline tagging of the corpora was carried out by Dr Anthony Hennessey.

output in order of probability. For my analysis with *CorporaCoCo*, I follow the procedure of *Wmatrix*, which selects only the first tag from the *USAS* output (P. Rayson, personal communication, September 24, 2017).

In general, the *USAS* precision rate has been reported as relatively high (91%; see Rayson, Archer, Piao, & McEnery, 2004). This value will depend on the corpus and might be slightly lower for the Surveillance Blog Corpus, which contains digital discourse. For example, the word *content* is often tagged as “E4.2+” tag in the sense of “content/satisfactory”, which forms part of the discussion on emotional tags in Section 5.4.2. Yet, in 34 of the 199 co-occurrences of the “M” tag (“Movement, location, travel & transport”) with “E4.2+” in the blog posts, the noun *content* exhibits an unemotional, digital content sense (as in *amateur *content* producers*) that may not have been incorporated into the *USAS* tagger yet. Regarding the thematic focus on surveillance, Section 5.4.1.1 reports that the abbreviation *NSA* is categorised as meaning “no strings attached”, although it refers to *National Security Agency* in the corpus. Balossi (2014) manually recategorises *USAS* tags that do not suit her data for her much smaller, specialised corpus consisting of one novel. To manually check all tags in the two semantically tagged corpora, amounting to over 5 million words, is beyond the scope of the present study. Overall, however, the results from the semantic tag comparison in Chapter 5 are useful, as demonstrated by the analysis of concordance lines in the context of examples.

3.2.5 Corpus-external data

The methodological framework of this thesis is based on a corpus linguistic approach that relies on textual data. However, it is undeniable that semiotic resources beyond written text have an impact on meaning in discourse and I recognise this potential in the analysis. Whereas it is beyond the scope of the present study to fully implement a ‘corpus-assisted multimodal

discourse analysis' framework (as introduced by Bednarek & Caple, 2017), I discuss multimodal aspects of the surveillance discourse as part of the analysis in Chapters 5 and 6. Here, I introduce the sources of the multimodal data; the procedural details of the analysis are given in the respective analysis chapter.

In Chapter 5, I put forward a corpus linguistic approach to the concept of surveillant landscapes, which derives from the heavily semiotically oriented tradition of mediated discourse analysis (see Sections 2.3, 2.6.1). To explain my corpus linguistic analysis of surveillant landscapes, I first carry out a case study (Section 5.2.1) following a conventional mediated discourse analysis approach to study the surveillant elements of a particular place, a local shopping centre (the Bullring in Birmingham, UK). I then complement that analysis with corpus linguistic findings for the same shopping centre in Section 5.2.2 and surveillant landscapes more generally in the rest of the Chapter. The data for the Bullring case study originate from a variety of sources, including a visit to the shopping centre in July 2018 and text and video content from its official website and YouTube channel. During the visit of the shopping centre I took photos of markers of surveillant elements, such as signs announcing the presence of CCTV cameras. Rather than offer an exhaustive study of the surveillant landscape of that mall, the case study is intended to illustrate how various multimodal elements interact in creating the surveillant landscape. Crucially, the aim of this case study is to show parallels between the mediated discourse analysis of a particular place and the corpus linguistic analysis on a larger scale. The argument is that co-occurrence patterns are central to both approaches: textual co-occurrence is an important concept in corpus linguistics and mediated discourse analysis can be said to analyse the way in which semiotic and material modes interact and co-occur. Section 5.2 explains these parallels in detail.

As Bednarek and Caple (2017) show in their book, multimodal meanings are especially important in the context of news media. In order to account for any semiotic resources appearing alongside the articles in *The Times*, I use the online interface of the TDA (GALE, 2018). The interface provides various functions to view the newspaper issues, including a “browse by date” function, a basic search with optional advanced settings as illustrated in Figure 3-4. As part of the advanced settings, it is possible to specify whether the search term should appear (e.g. in “entire documents” or just in the title) and to restrict the search to particular publication dates and/or publication sections. If no particular section is chosen, the results screen includes an overview of the distribution of documents across the sections. I use this function to identify the newspaper sections in which the search terms occur most frequently for my analysis in Section 6.4. Figure 3-5 shows an example of the output from the search for *surveillance* in 2008, with the distribution of documents containing the word across publication sections displayed in the bottom left corner (“News”, “Arts and Sports”, “Business News”, “Opinion and Editorial”, “Advertising”, “People” and “Preliminary and Supplementary Material”).

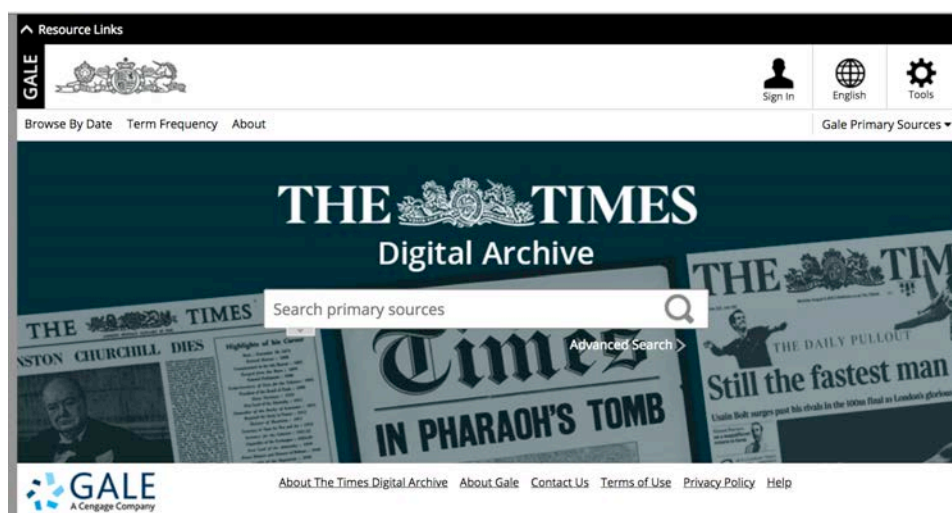


Figure 3-4: The Times Digital Archive online interface

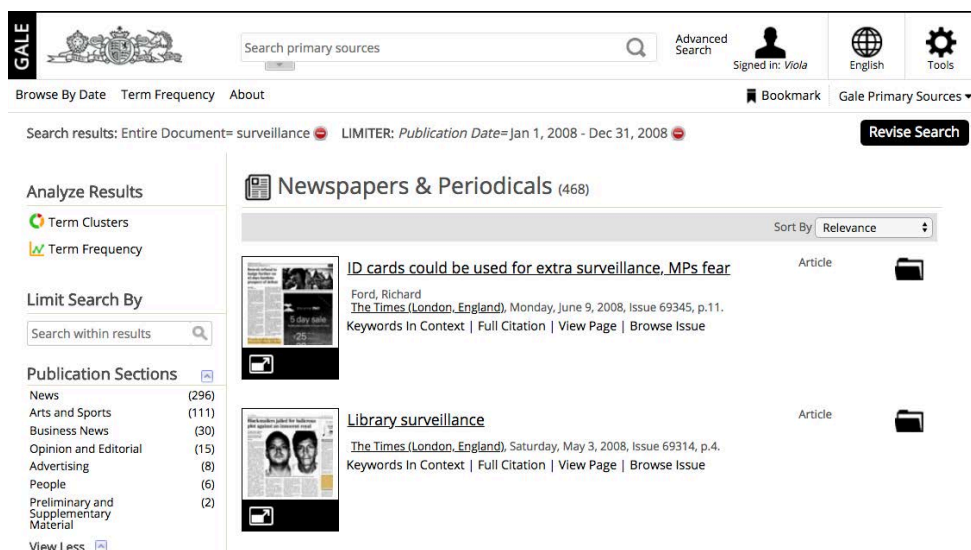


Figure 3-5: TDA advanced search for *surveillance* in 2008, showing the distribution across publication sections

3.3 Analysis

The three analysis chapters of this thesis fundamentally follow the same procedure consisting of a co-occurrence comparison and contextualising the identified co-occurrence patterns through concordance analysis. Section 3.3.1 outlines the steps of the co-occurrence comparison method. Due to the different natures of both the research questions and the corpora employed, the implementation of this procedure differs for each chapter. These customised approaches are explained in Section 3.3.2, respectively. Finally, Section 3.3.3 outlines some limitations of the co-occurrence comparison method.

3.3.1 Co-occurrence comparisons

This section introduces the methodological framework of co-occurrence comparisons, which is broadly followed in all analysis chapters. The framework applies the pairwise co-occurrence comparison method that colleagues and I have developed (see Wiegand et al., 2017b). The comparison is carried out with the corresponding algorithm implemented in the *CorporaCoCo* R package. This thesis has acted as a testing ground for the functionality and development of the package. As such, the stages of the analysis make use of different versions of the package

and reflect the development of the research undertaken in this thesis as well as the methodology of co-occurrence comparisons. So I used an unpublished development version for Chapter 4 (version 0.13) and a development version of the published release version 1.1 (Hennessey et al., 2017).¹⁴

The flowchart in Figure 3-6 outlines the co-occurrence comparison procedure in *CorporaCoCo*. The starting point is the research question. This question is crucial for step 2, the selection of the two sets of data that are required for the comparison: the two corpora to be compared (“Corpus A” and “Corpus B”) and a set of node words for which the co-occurrences are counted and compared. It would not be good practice to use all types in Corpus A and B as the node words, because, by random chance, a large number of false positives would be found.

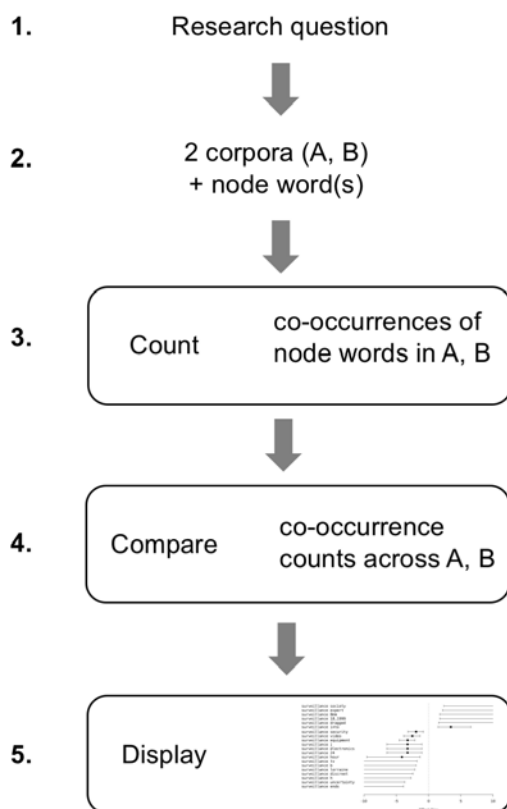


Figure 3-6: Flowchart of the co-occurrence comparison method using *CorporaCoCo*

¹⁴ *CorporaCoCo* runs in an R environment (R Core Team, 2016). I also used *RStudio* (RStudio Team, 2015) for my analysis.

For Chapter 4, which used Version 0.13 of *CorporaCoCo*, I imported plain text files into *R* and tokenised them with a separate text processing package, choosing the *quanteda* package at that stage in my research (Benoit & Nulty, 2016) to save the tokenised texts in vectors for use in *CorporaCoCo*. With the 1.1 development version of *CorporaCoCo*, the import process has been formalised – *CorporaCoCo* itself employs the tokenisation from the *stringi* package (Gagolewski et al., 2017). The text is saved in a ‘corp_text object’, an *R* list object that “encapsulates the tokenization of a piece of text” (Hennessey et al., 2017). This newer version of the package was used for both Chapters 5 and 6. However, only the analysis of the TDA1986–2008 in Chapter 6 is based on the Unicode tokenisation (see Section 3.2.3.2), because the semantically tagged corpora in Chapter 5 are tokenised according to the *CLAWS* tagger (see Section 3.2.4).

Figures 3-7 and 3-8 show sample outputs for extracts from corp_text objects from the TDA1986–2008 and the Surveillance Blog Corpus, respectively.¹⁵ They illustrate the basic concept of the corp_text object, showing the running number of the token in the corpus file (“idx”), the type (in Figure 3-7 a simple lower case version of the token; in Figure 3-8 semantic tags), a start and end counter to indicate the position in the file and, finally, the tokens as found in the original text. For the Surveillance Blog Corpus, I make use of the capability of *CorporaCoCo* to deal with ‘complex types’, i.e. instead of the lower case tokens we now assign the semantic tags to the types column of the corp_text object.

¹⁵ I produced this output using the *CorporaCoCo* `corp_get_tokens()` function.

idx	type	start	end	token
1: 32116	anxious	185891	185897	Anxious
2: 32117	that	185899	185902	that
3: 32118	they	185904	185907	they
4: 32119	may	185909	185911	may
5: 32120	be	185913	185914	be
6: 32121	left	185916	185919	left
7: 32122	with	185921	185924	with
8: 32123	too	185926	185928	too
9: 32124	much	185930	185933	much
10: 32125	merchandise	185935	185945	merchandise
11: 32126	retailers	185948	185956	retailers
12: 32127	have	185958	185961	have

Figure 3-7: Sample output of a TDA corp_text object (extract from monthly_object2008_12.rds)

idx	type	start	end	token
1: 243	Q2.1 Q1.2	1333	1338	Noting
2: 244	A6.1-	1340	1344	other
3: 245	T2- N4	1346	1351	former
4: 246	W1 S5+c A4.1 N5+	1353	1357	world
5: 247	S7.1+/S2mf A5.1+++ /S2mf Q4.2 G2.1/S2mf	1359	1365	leaders
6: 248	Z1mf Z2	1368	1371	Lyon
7: 249	M1 X2.4	1373	1380	explores
8: 250	Z5	1382	1384	the
9: 251	T1.1.1 T3+	1386	1395	historical
10: 252	N3.2+/A2.1 A5.1+/A2.1 L3 B2-	1397	1402	growth
11: 253	Z5	1404	1405	of
12: 254	X2.4	1407	1418	surveillance
13: 255	Z5	1420	1423	from
14: 256	A13.6 A14 A12+	1425	1430	simply
15: 257	A9+ N6+ T2++ A1.7+ A2.2 F4 S8+ H4	1432	1438	keeping
16: 258	O2	1440	1443	tabs
17: 259	Z5	1445	1446	to
18: 260	Z5	1448	1450	the
19: 261	T3-	1452	1454	new
20: 262	X4.1	1456	1462	concept
21: 263	Z5	1464	1465	of
22: 264	N3.2+ N5+ A11.1+ S1.2.5+ X5.2+	1469	1471	Big
23: 265	S4m S2.2m S9/S2.2m	1473	1479	Brother
24: 266	Z5	1483	1485	and
25: 267	Z8	1487	1489	its
26: 268	A1.2+	1491	1499	relevance
27: 269	Z5	1501	1502	in
28: 270	T3-	1504	1509	modern
29: 271	S5+c	1511	1517	society

Figure 3-8: Sample output of a Surveillance Blog Corpus corp_text object (extract from blogspot_outfile056.rds)

In the third step, the raw co-occurrences for each node are counted in each corpus. This thesis uses surface co-occurrence, because that is the traditionally most wide-spread co-occurrence

type in corpus linguistics. As explained in Section 2.4.4, for this co-occurrence type, collocates are identified in a given span. All co-occurrence comparisons in this thesis use a span of five words to the left and right, which is a commonly used span and therefore a reasonable choice unless the research question calls for a different span.

Figure 3-9 illustrates the processes of co-occurrence counting for the node *surveillance* and the collocate *security* in two short sample corpora. Both corpora contain three instances of the node, with the span highlighted in colour and the collocate marked by asterisks. In the counting step, *CorporaCoCo* goes through each span in the corpora to count a ‘hit’ when a span contains an instance of the collocate and a ‘miss’ for each slot in a span that does not contain the collocate. In Corpus A, we find that each instance of *surveillance* co-occurs with *security* in the given span, adding up to three hits for the co-occurrence pair (*surveillance*, *security*) in Corpus A. By contrast, Corpus B only contains one hit of the collocate *security* (in the “red span”). The number of misses is calculated by adding up all the coloured token slots that do not contain *security*.

[CORPUS A: extract from 2012-09-3-09 in the S&S Corpus]

As such, it contributes a socio-legal perspective to the literature on *security* and **surveillance** practices that has until now focused chiefly on *security* professionals, policing, military, **surveillance** and intelligence techniques. It therefore takes seriously the notion that *security* and **surveillance** are techniques of governing that need to be analysed not for their own sake, but in relation to a broader setting.

hits: 3 x 1 = 3

misses: 3 x 9 = 27

rate of co-occurrence: $\frac{hits}{misses} = \frac{3}{27} \approx 0.111$

[CORPUS B: extract from 2004-02-2_3-13 in the S&S Corpus]

The public notion of video **surveillance** as a means to prevent crime seems to meet that goal best. At the same time, management does not want customers to know that they are under **surveillance** because they think this could inhibit potential consumers from feeling at ease in the mall. The technical or security managers mostly agree with the centre managers. They more often cite objectives of property management: e.g. guarantee of workflow, cleanliness and monitoring the function of technical devices (e.g. elevators, doors, escalators). One *security* manager admitted that video **surveillance** can also be used to reduce staff.

hits: 1

misses: 2 x 10 + 1 x 9 = 29

rate of co-occurrence: $\frac{hits}{misses} = \frac{1}{29} \approx 0.034$

	Corpus A	Corpus B	
Hits	3	1	4
Misses	27	29	56
	30	30	60

Figure 3-9: Illustration of the co-occurrence comparison procedure¹⁶

The co-occurrence counts in the format of hits and misses provide the basis for the comparison (step 4). By dividing the number of hits and misses we can derive co-occurrence rates for the co-occurrence pair in each corpus: in the example from Figure 3-9 the co-occurrence rate for

¹⁶ This example is adapted from an example that colleagues and I showed with different texts in Wiegand et al. (2017b).

surveillance and *security* is 0.111 in Corpus A and 0.034 in Corpus B. These rates can give an indication of how similarly or differently the words co-occur in the corpora. Yet to statistically determine whether the co-occurrence rates are significantly different, a significance test has to be carried out. *CorporaCoCo* runs a Fisher’s Exact Test of the hits and misses in the two corpora, which are represented in the contingency table in Figure 3-9.¹⁷

The results of the co-occurrence comparisons are co-occurrence pairs that are significantly more frequent in either Corpus A or Corpus B. In step 5 of the methodology, these results are displayed. Figure 3-10 gives an example plot (analysed in Section 5.3.1.1) that compares co-occurrences of two different nodes, *space* and *place*, in the Surveillance Blog Corpus (left) and the S&S Corpus (right). Each co-occurrence pair represents a separate comparison and the black box next to the co-occurrence pair indicates the effect size, which is a value for the extent to which the co-occurrence counts differ between the two corpora. When the effect size box is missing, this means that the co-occurrence pair is “unique” in that corpus – i.e. it does not appear in the other corpus – and therefore the effect size cannot be calculated. The whiskers on which the effect sizes boxes are placed represent the confidence intervals, indicating how much evidence there is for each result. The shorter the whiskers, the more evidence there is for the difference.

¹⁷ For this small-scale example, chosen for illustrative purposes, the test does not find a significant difference: p-value = 0.6119536; confidence-interval (95%): (-7.45, 2.07).

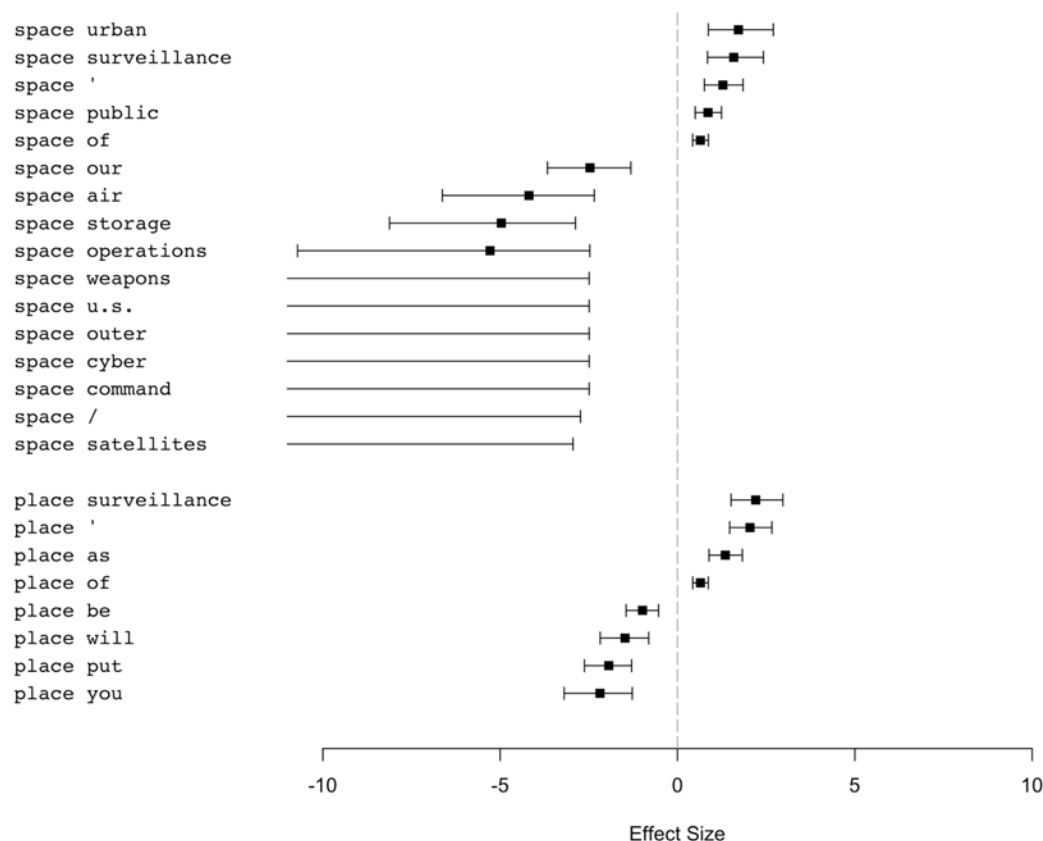


Figure 3-10: Sample *CorporaCoCo* plot

As with other types of quantitative corpus linguistic results, the significantly different co-occurrence pairs form the starting point for the analysis rather than an end result in themselves. I analyse concordance lines of the node with the identified co-occurrences to analyse the function of the co-occurrence pairs in context. The concordance analysis is closely related to the respective theoretical focus of the chapter.

As each co-occurrence pair represents an individual comparison, a large number of tests are carried out even when only a small number of node words are chosen. The larger the number of tests, the larger the chance that false results will be included. Statisticians deal with this problem with ‘multiple testing corrections’. *CorporaCoCo* implements one of these techniques, which is called the ‘false discovery rate’ (FDR; Benjamini & Hochberg, 1995). I follow the recommended default FDR value of 0.01 implemented in the package, i.e. only one in 100

results is expected to be false by random chance. In Chapter 6 only, I increase the FDR value to 0.1 so as not to curb the small number of results too strictly. With this setting, up to one in ten results can be caused by random chance. Crucially, I follow up the co-occurrence comparisons with concordance analyses and searches in the online interface of the TDA.

By default, the comparison counts the co-occurrences of the chosen node words with all collocates found in the given span of Corpus A and Corpus B. If the research question can be narrowed down to particular co-occurrence pairs, the efficiency of the test can be further improved. So, an optional setting allows the user to pick particular collocates to be counted and compared rather than comparing all collocates. I make use of this option in Chapter 5, where for part of the analysis (in Section 5.4) I focus on collocates from the semantic fields of social aspects and emotion, as identified with the *USAS* semantic tagger.

3.3.2 Implementing co-occurrence comparisons in individual chapters

The methodology of co-occurrence comparisons forms the core of this thesis. This section explains the implementations of the comparison in the individual chapters. It only aims to provide a broad overview of how the co-occurrence comparison methodology is implemented for the various studies. Further details of the procedures within the context of the chapter's research questions are then given in Chapters 4–6, respectively.

Table 3-10: Corpus A/B and node words for co-occurrence comparisons Chapters 4–6

	Corpus A	Corpus B	Nodes
Chapter 4	1 volume subcorpus of the S&S Corpus	S&S Corpus	KKWs of the S&S Corpus
Chapter 5	Surveillance Blog Corpus	S&S Corpus	Lexical and semantic sets of “place” nodes
Chapter 6	A & B: Yearly & monthly subcorpora of the TDA		<i>surveillance, privacy, CCTV</i> ; identity card-related nodes

Table 3-10 provides an overview of the selection of Corpus A and B and node words for each analysis chapter. As the first analysis chapter, Chapter 4 acts as an introduction to surveillance discourse for this thesis. The chapter traces the developments across the discourse of the specialist academic S&S journal and starts with an analysis of surveillance definitions and identifying words that are consistently salient across all 13 volumes of the journals: key keywords. The key keyword (KKW) analysis involves determining keywords for each individual volume. For this keyword comparison, the (original) British National Corpus (BNC; Aston & Burnard, 1997) is used as a reference corpus. The BNC is a widely used reference corpus and popular for its considerable size, representation and public availability. While the texts in the S&S Corpus were published more recently than the BNC text – most of which were published between 1975–1993 (see Aston & Burnard, 1997, p. 30), there is only a gap of one decade between the latest BNC texts (1993) and the earliest S&S texts (2002). At the time of writing, the written component of the new “BNC2014” Corpus, which would be an even better fit for the time period of the S&S Corpus, is not yet publicly available (see ESRC Centre for Corpus Approaches to Social Science (CASS), n.d. for updates on the project).

The set of keywords that are key in each volume qualify as KKWs. Section 4.3 explains the procedure for identifying the KKWs (using *WordSmith Tools*) and discusses them in terms of thematic categories. The premise is that these KKWs point to central concepts of the corpus.

The co-occurrence comparison is then used to investigate how these KKWs differ in their usage across the corpus. Pairwise comparisons are carried out for each volume compared against the full corpus, i.e. each of the 13 journal volumes in turn takes the position of Corpus A to be compared against the full S&S Corpus as Corpus B. The most frequent KKWs are taken as nodes for the co-occurrence comparisons in Section 4.4, which are analysed in relation to the themes that Mehrabov (2015) has identified in the S&S Journal. Whereas all 69 KKWs

could theoretically function as nodes, this would produce a large number of potentially misleading and irrelevant co-occurrences. In the final analysis stage of Chapter 4, I present a network of shared patterns across the salient collocates of *surveillance*.

Chapter 5 compares the co-occurrence patterns of place references in the Surveillance Blog Corpus with the academic S&S corpus as the reference corpus. In this chapter, the main analytical focus is on place patterns in relation to the concept of the surveillant landscape (see Section 2.6.1). As indicated in Table 3-10, Corpus A is the Surveillance Blog Corpus and it is compared against the S&S Corpus from Chapter 4 as Corpus B, which acts as the reference corpus here. To facilitate the analysis of the surveillant landscape, nodes related to place and mobility are chosen. The comparison starts with several hand-picked place words for comparing lexical co-occurrences (Section 5.3.1), in a procedure akin to that from Chapter 4. However, the analysis then moves on to a larger-scale comparison that involves all words belonging to the semantic field related to places or mobility (5.3.2 and 5.4). This comparison is based on the semantic category of “Movement, location, travel and transport” in the *USAS* tagset (out of the 21 “major discourse fields”; Archer et al., 2002, p. 2). The chapter therefore expands the concept of co-occurrence comparisons from single words to semantic fields.

If the complexity of Chapter 5 is in its selection of nodes, then the contribution of Chapter 6 is its development of a systematic method of comparing subcorpora of a newspaper corpus to investigate diachronic co-occurrence change. A simpler version of this arrangement of pairwise comparisons is already used in the comparison of journal volumes Chapter 4, but it is the temporal structure of the TDA1986–2008 that allows for the actual diachronic comparison. Chapter 6 begins with pairwise “long-term” comparisons across the first, middle and final year of the corpus (1986, 1997, 2008) to investigate overall co-occurrence trends. These are discussed in Section 6.2.

The chapter then moves on to “peak comparisons” where Corpus A represents a monthly period with peak relative frequency of one of the node words. The node words for this part of the analysis are *surveillance*, *privacy* and *CCTV* to represent different aspects of the surveillance discourse. For each node word, the top five relative frequency peak months are identified and then employed as Corpus A in a comparison. For these comparisons, the 1986 subcorpus is used as Corpus B to function as a baseline. Only for the node *CCTV*, the year 1999 is used as a baseline instead, owing to the low frequency of this node in the early years of the corpus. The peak comparison procedure is explained in Section 6.3.

Chapter 6 contains an additional case study that also makes use of the *CorporaCoCo* analysis, framed by the public debate about the UK Identity Cards Act 2006. For this analysis, the year 2002 is taken as a baseline and the following years are compared against it. The nodes for this analysis relate to the topic of identity cards (which were originally proposed as “entitlement cards”): *identity*, *id*, *entitlement*, *database*, *card(s)*. This comparison is explained in Section 6.5.

In addition to the co-occurrence comparisons, each analysis chapter contains further analysis techniques that are only reported in the chapters. These include key keyword analysis (Chapter 4), the mediated discourse analysis interpretation of the Bullring shopping centre (Chapter 5) and the analysis of distribution patterns across newspaper sections (Chapter 6).

3.3.3 Limitations of co-occurrence comparisons

I have been working with various development versions of *CorporaCoCo*, which has facilitated an innovative methodology. Yet, because the method has been developed recently and this thesis presents the first large-scale application of *CorporaCoCo*, some aspects are still under development. Dealing with different versions of the package as well as different datasets has

complicated the methodology of this thesis and meant that, for example, different types of tokenisation have been used for the different chapters. However, tokenisation has been consistent for individual comparisons (e.g. the versions of the Surveillance Blog Corpus and the S&S Corpus studied in Chapter 5 are both tokenised by the *CLAWS* tagger).

As of *CorporaCoCo* 1.1 (development version), the package does not support the analysis of multiword units. Handling multiword units would be a useful feature for studying collocates of compound words like *surveillance society* and *identity cards*, for example. A potential solution would be to use a workaround, such as merging all instances of *surveillance society* in the corpus. This approach would have been most helpful for the analysis of the TDA data in Chapter 6, where these compounds are discussed in most detail. I have not merged any compounds in the TDA. Given the large size TDA1986–2008 (and the time it took to process the `corp_text` objects), it was not feasible to create multiple versions of the corpus for this study.

Another challenge I have faced is a common issue in corpus linguistics: the comparisons produce large numbers of results (even when multiple test corrections is used). This challenge is amplified in some of the analysis stages of Chapter 5 related to semantic tagging. So, the analysis of the collocates of the major discourse domain “M” (“Movement, location, travel & transport”) in Section 5.3.2.2 produces so many results that the plot cannot be printed. I therefore restrict my discussion to the tags with the highest effect size. Given the number of comparisons made in this thesis, it was not possible to analyse the concordances of all collocates that were identified. So, I have tried to focus on collocates that are either salient because of (i) their large effect sizes and narrow confidence intervals or (ii) theoretical reasons. Finally, it must be said that the *CorporaCoCo* method highlights statistical differences in recurrent textual patterns. I have tried to examine meaning beyond these results by looking for similarities across the patterns (see the network of surveillance collocates in Section 4.5), exploring additional,

multimodal data sources (see Section 3.2.5) and examining examples in context throughout the analysis chapters.

3.4 Conclusions

This chapter has presented the methodological framework of this thesis, introducing the corpora and the main method of analysis. In following the three principles of meaning-making, the methodological framework investigates (i) naturally occurring language (ii) comparatively across different discourse domains and (iii) focuses on co-occurrence patterns. While the main emphasis is placed on lexical co-occurrence patterns, this approach is extended to semantic patterns in Chapter 5. In addition, I qualitatively examine multimodal meaning-making patterns that form links between textual elements and other semiotic modes in Chapters 5 and 6 via the mediated discourse analysis of the surveillant landscape and visual resources in the TDA.

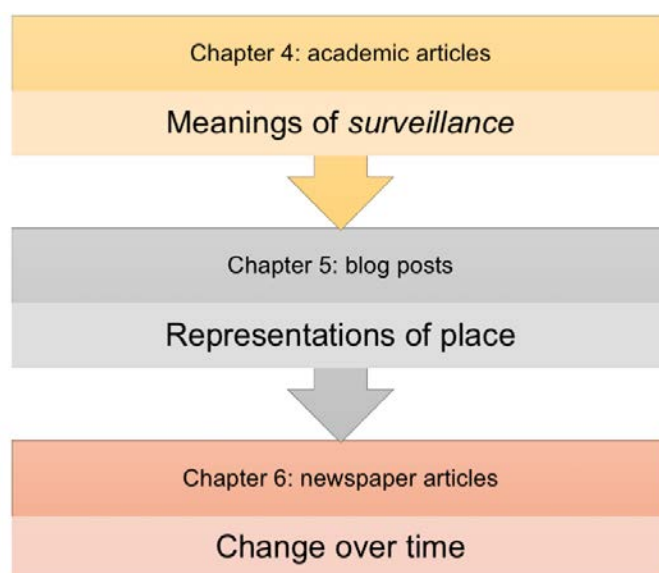


Figure 3-11: Methodological workflow for analysing surveillance discourse

So, the methodological approach taken to surveillance discourse in this thesis can be summarised as sketched out in Figure 3-11. The analysis begins in Chapter 4 with a study of

the meanings of *surveillance* found in “expert” texts by surveillance studies scholars. The thesis then moves on to patterns related to the two fundamental coordinates of discourse, place and time (see Section 2.6) in Chapters 5 and 6, by analysing the representation of surveillant landscapes and diachronic meaning change.

4 The academic representation of the concept of surveillance in the *Surveillance & Society* journal

4.1 Introduction

This chapter presents the beginning of the main analysis, focusing on the first of three domains of public discourse: academic writing. The *Surveillance and Society* (S&S) journal is an important platform for informing research and discussions on surveillance theory in the interdisciplinary field of surveillance studies (see Section 3.2.1). Because the S&S Corpus is formed of texts written by surveillance specialists, it is a good starting point for the present study of surveillance discourse. The methodology incorporates the meaning-making principles outlined in Chapter 2 that meaning evolves with the discourse, emerges via comparison and takes shape in co-occurrence patterns. Due to these principles, the S&S Corpus is not assumed to be a static, homogeneous representation of meaning. Instead, the chapter traces and compares co-occurrence patterns across a 14-year publication period. The assumption is that the analysis will find shifts, however subtle, in the representation of surveillance across this period. To avoid a “difference bias”, the analysis also looks for shared patterns.

The chapter approaches the discursive representation of surveillance in the S&S Corpus in three steps. The analysis begins with linguistic patterns for defining *surveillance* across the entire corpus based on frequent clusters and a qualitative concordance analysis of explicit definitions (Section 4.2). This stage of the analysis answers RQ 1-1 (“How is the concept of surveillance defined in the S&S journal?”). In the second stage (Section 4.3), KKWs of the S&S Corpus are identified to indicate salient and consistent lexical constituents of the surveillance discourse across the corpus. This stage answers RQ 1-2 (“Which words are

consistently salient across the S&S journal volumes?”). The co-occurrences of the most frequent KKWs, including *surveillance*, are compared across the journal volumes to uncover meaning shifts in their usage (Section 4.4). The discussion of these co-occurrence results is framed according to the main themes that Mehrabov (2015) has identified in special issues of the S&S Journal. Section 4.5 rounds off the analysis with a focus on wider patterns in the surveillance discourse of the journal presented with a network of salient collocates of *surveillance*. Together, Sections 4.4 and 4.5 address RQ 1-3 (“How do the meanings of the consistently salient words shift across the S&S journal volumes?”). Finally, Section 4.6 concludes the chapter.

4.2 Defining *surveillance*

Surveillance is very frequent in the S&S Corpus: it is among the top 20 most frequent words in the corpus, as shown in Table 4-1. The remaining top ranks are function words like determiners and prepositions or forms of BE. With a frequency of 20,265, *surveillance* is, while considerably less frequent than the top ranking function words like *the* and *of*, even more frequent than the pronoun *this* – and all personal pronouns that do not make it into the top 20. The high frequency of *surveillance* suggests that (i) it is used in many different contexts and, therefore, (ii) its meaning is likely to encompass a variety of components.

Table 4-1: The 20 most frequent words in the S&S Corpus (# stands for all numbers)¹

Rank	Word	Freq.
1	the	166,081
2	of	114,146
3	and	88,094
4	to	70,900
5	in	58,124
6	a	50,331
7	#	47,884
8	that	34,247
9	is	33,803
10	as	27,949
11	for	21,493
12	surveillance	20,265
13	this	18,180
14	are	18,072
15	on	16,971
16	by	16,098
17	it	15,597
18	with	15,538
19	be	15,317
20	or	14,069

To identify typical patterns of *surveillance* and associated meanings, we can examine the most frequent clusters formed with *surveillance*. Table 4-2 lists three-word clusters. The top three clusters are combinations of *surveillance* with highly frequent function words (see Table 4-1): *the*, *of*, *and*, *in*. On the following ranks we find content words that provide more information on the meaning of *surveillance*. In Section 4.2.1, I examine thematic groups of frequent clusters for an initial overview of meaning constituents of *surveillance*. Section 4.2.2 then provides a detailed concordance analysis of explicit definitions of *surveillance* by the surveillance studies scholars in the S&S Corpus.

¹ All frequencies reported in this chapter are based on *WordSmith Tools*. Only the co-occurrence comparison plots in Section 4.4 have been created with *CorporaCoCo*.

Table 4-2: The 20 most frequent three-word clusters in the *surveillance* concordance (using the *WordSmith Tools* Version 7.0.0.167 Index Cluster function “running clusters from a selection” for *surveillance*, “length 3, maximum working percent 100%, stop at sentence break”)

Rank	Cluster	Freq.
1	of surveillance and	414
2	of the surveillance	275
3	of surveillance in	263
4	forms of surveillance	254
5	the surveillance of	225
6	of surveillance studies	215
7	surveillance in the	209
8	of surveillance technologies	193
9	surveillance and the	156
10	in surveillance studies	149
11	surveillance as a	149
12	of surveillance as	140
13	of video surveillance	139
14	of surveillance is	132
15	use of surveillance	127
16	of surveillance that	121
17	public health surveillance	112
18	form of surveillance	102
19	surveillance and control	99
20	in the surveillance	98

4.2.1 Clusters and meaning components of *surveillance*

For this initial overview I highlight three groups of *surveillance* clusters that

- i. point to the diversity of the concept – e.g. *form(s) of surveillance*;
- ii. describe the academic discipline – e.g. *of/in surveillance studies*; and
- iii. contain a form of BE – e.g. *of surveillance is*; and, further down the cluster list (beyond Table 4-2), *surveillance is a* and *surveillance is not*.

(i) Clusters pointing to the diversity of the concept

The cluster on rank four, *forms of surveillance*, explicitly points to multiple manifestations of the concept and therefore supports the impression that *surveillance* is used in many contexts.

Concordance 4-1 shows examples that particularly emphasise this diversity: *diverse*, *many*, *multiple* and *various forms of surveillance*. Two additional and more frequent patterns are not included for reasons of space: *different forms of surveillance* occurs ten times and *other forms of surveillance* 20 times. In addition, we find examples indicating that this diversity is not static, but in development: *emerging* (three occurrences), *contemporary* (six occurrences) and *new(er) forms of surveillance* (21 occurrences) in contrast to *traditional* (three occurrences), *historical* (two occurrences) and *old forms of surveillance* (one occurrence). The full concordance distinguishes between *forms of surveillance* based on their characteristics (e.g. *bureaucratic*, *centralised*, *decentralised*, *mobile*, *non-technological*). Finally, we find mentions of forms alluding to controversy and ethical questions: *coercive*, *contentious* and *egregious* (see Concordance 4-2).

1 nd submitting to diverse forms of surveillance is normal, necessary, a 2014-12-3-07
2 t years the most diverse forms of surveillance have been found at airp 2003-01-4-04
3 eillance practices. Many forms of surveillance have been introduced to 2014-11-4-01
4 e the objects of so many forms of surveillance aimed at controlling th 2014-11-4-01
5 tations inherent in many forms of surveillance, as well as in the proc 2013-11-3-04
6 not overcoded with many forms of surveillance potential, as has been 2012-10-3-4-01
7 articipating in multiple forms of surveillance. However, because of it 2014-12-3-07
8 es subjected to multiple forms of surveillance, coercion and violence 2010-08-2-16
9 o to resist the multiple forms of surveillance (Lyon, 2001). In one se 2004-02-2-3-06
10 acknowledge the multiple forms of surveillance that occur in many clas 2003-01-4-05
11 o to resist the multiple forms of surveillance. However, resistance ma 2003-01-3-05
12 eworks. However, various forms of surveillance permeate ordinary every 2014-12-4-09
13 als to engage in various forms of surveillance (Foucault 1979) that in 2014-11-4-11
14 national sport, various forms of surveillance are part of the routine 2014-11-4-01
15 associated with various forms of surveillance have largely engaged wi 2010-08-2-08
16 y's concern that various forms of surveillance might have discriminato 2009-06-4-07
17 es feed into the various forms of surveillance that continue to prolif 2005-03-2-3-09

Concordance 4-1: Examples of *forms of surveillance* (17 out of 254) – diversity

1 rder' or more 'coercive' forms of surveillance including having drinks 2010-07-3-4-06
2 rder" or more "coercive" forms of surveillance, such as being escorted 2010-07-3-4-01
3 eed, various contentious forms of surveillance are intrinsically conne 2014-11-4-01
4 , these more contentious forms of surveillance are too often eclipsed 2010-08-2-07
5 acilitate more contested forms of surveillance (Humphreys 2008, 2009; 2014-12-3-02
6 he presence of egregious forms of surveillance. To this end, in this p 2011-09-1-2-01
7 he presence of egregious forms of surveillance. In this paper, we disc 2011-09-1-2-01
8 ies in combating extreme forms of surveillance as a property managemen 2004-02-4-04
9 scale and more invasive forms of surveillance appear to us as the nec 2011-08-4-01
10 lematic and unacceptable forms of surveillance, which depended on thei 2012-10-3-4-02

Concordance 4-2: Examples of *forms of surveillance* (10 out of 254) – controversy

- (1) [...] CCTV is put in public space by governmental or private sector agencies, thus representing a top-down **form of surveillance** [...] [2012-10-2-05; (Timan & Oudshoorn, 2012)]

The singular variation, *form of surveillance*, though not quite as frequent as the plural one, also appears in Table 4-2. Example (1) illustrates how this cluster tends to refer to a meaning component of *surveillance* relevant to the given context, positioning CCTV technology as “a top-down form of surveillance” (later contrasted with mobile cameras which represent a “bottom-up form”). Like technological or domain-related *forms*, the example supports the impression that *surveillance* is depicted as a concept that is found in many domains.

(ii) *Clusters describing the academic discipline*

The second group of clusters directly refers to the academic discipline. In Table 4-2, these clusters include *of surveillance studies* and *in surveillance studies*. Further down the list is the cluster *field of surveillance* (rank 27, 84 occurrences) that overlaps with the first cluster, *of surveillance studies*, which I focus on here.

Concordance 4-3 shows that *(the field) of surveillance studies* is often described as *inter-*, *cross-* or *transdisciplinary* (lines 3–6, 9–11, 13, 17 – where *interdisciplinary* is only partly shown, and 19). Research can be situated at the “intersection of surveillance studies” with other fields (lines 14–16; also see Lyon, 2007). While not using these terms, additional lines relate *surveillance* to *diversity* (line 1), other *diverse fields* (line 12), and portray it as a *vast field* (line 2) or a *broad tradition* (line 18). These descriptions echo the journal’s editorial “focus and scope” statement, which specifically mentions “transdisciplinary work on surveillance” among its publication aims (“About the Journal,” 2019, online; see Section 3.2.1). Surveillance studies crosses not only disciplinary but also national boundaries (*transnationalization*; line 20). The

combination of different scholarly approaches is a response to the complexity of the concept of surveillance (as observed in Concordances 4-1 and 4-2).

1	illustrate the diversity of Surveillance Studies and provide a useful r	2009-07-1-08
2	Mapping the vast field of Surveillance Studies is not easy, since Sur	2015-13-1-09
3	multidisciplinary field of Surveillance Studies since the 1980s, might	2015-13-1-04
4	interdisciplinary field of Surveillance Studies have described dangero	2014-12-2-04
5	interdisciplinary field of surveillance studies itself, these works dr	2012-09-3-11
6	ross-disciplinary field of Surveillance Studies. Besides police state	2011-09-1_2-06
7	rous and shifting field of Surveillance Studies.	2009-06-3-16
8	hose in the wider field of surveillance studies. While the theoretical	2009-06-3-16
9	interdisciplinary field of surveillance studies usefully takes up the	2008-05-3-02
10	transdisciplinary field of surveillance studies suffers from an overab	2005-03-2_3-03
11	interdisciplinary field of surveillance studies. The social and materi	2005-03-1-01
12	from the diverse fields of Surveillance Studies, kinesiology, sports p	2014-11-4-01
13	ransdisciplinary flight of surveillance studies has been usefully augm	2009-06-3-16
14	tives. The intersection of Surveillance Studies and IR has already pro	2015-13-2-04
15	s from the intersection of Surveillance Studies and his technical trai	2015-13-2-03
16	linary. An intersection of Surveillance Studies, political philosophy,	2015-13-2-03
17	disciplinary literature of surveillance studies acknowledges that surv	2010-08-1-03
18	he very broad tradition of surveillance studies (Murakami Wood 2009).	2012-09-4-01
19	and transdisciplinarity of Surveillance Studies as I know it. Perhaps	2009-06-1-08
20	he transnationalization of surveillance studies, the authors present t	2008-05-2-01

Concordance 4-3: Examples indicating interdisciplinarity and diversity (20 out of 215 occurrences of *of surveillance studies*)

Another theme in the concordance lines for *of surveillance studies* is the “development” of the field, as illustrated in Concordance 4-4 (particularly lines 2–8). Overall, the impression is that surveillance studies is *growing* (see lines 10, 13, 15, 16), has a *growing speed* (line 20), and is *strengthening* (line 21). The field is *exciting* (line 1) and *burgeoning* (lines 12, 14). Yet, the development is not equally strong in all aspects and regions. Concordance 4-4 prominently foregrounds the role of surveillance studies in Latin America, which is the theme of special issue 10(1). Historically, the development of the discipline in that geographical area is described as “hampered” or “limited” (lines 2–3) and its “growing speed [...] is variable” (line 20). So, the special issue sets out to contribute towards the development of surveillance studies in this region.

1 the most exciting area of Surveillance Studies. Both Hier and Greenbe 2009-06-1-08
2 the limited development of surveillance studies in Latin America, it m 2012-10-1-02
3 ampered the development of surveillance studies in Latin America. To d 2012-10-1-02
4 tion in the development of Surveillance Studies. I build on the work o 2011-08-4-13
5 , the rapid development of surveillance studies in the mid-2000s can a 2009-07-1-04
6 tial to the development of Surveillance Studies. The late 1980s and ea 2009-06-1-08
7 tion in the development of Surveillance Studies that goes beyond curre 2005-03-2_3-07
8 tion in the development of Surveillance Studies would be to embrace th 2005-03-2_3-07
9 he historical evolution of Surveillance Studies and its interactions w 2015-13-1-09
10 d rapidly growing field of Surveillance Studies to attend to gender (B 2015-13-1-04
11 urprisingly - the field of Surveillance Studies has developed in ways 2014-12-2-04
12 rgeoning academic field of Surveillance Studies, covering most if not 2012-10-3_4-17
13 ed in the growing field of surveillance studies, allows us to see how 2011-08-4-03
14 of the burgeoning field of surveillance studies. Drawing on establishe 2009-07-1-11
15 on...to the growing field of surveillance studies." We also agree with A 2009-06-4-07
16 ut to the growing field of surveillance studies as well. The report bu 2009-06-3-11
17 pidly developing field of 'Surveillance Studies' has sparked a remarka 2009-06-1-04
18 political significance of surveillance studies. By encouraging an ind 2009-07-1-07
19 oretical sophistication of Surveillance Studies with lawyerly attentio 2015-13-1-07
20 ever, the growing speed of surveillance studies in Latin America is va 2012-10-1-01
21 nt in the strengthening of surveillance studies on public security, ma 2012-10-1-02

Concordance 4-4: Examples indicating interdisciplinarity and diversity (21 out of 215 total occurrences of *of surveillance studies*)

(iii) *Clusters containing a form of BE*

The third group of clusters include a form of BE. Table 4-2 contains the form *of surveillance is*.

However, I focus on two clusters beyond the top ranks: *surveillance is not* (79 occurrences) and *surveillance is a* (75 occurrences). Concordances of these two clusters provide a particularly good place for the paraphrases that constitute the meaning of a discourse object (see Teubert, 2010).

I begin with the cluster that describes the characteristics of *surveillance* rather than the one that negates them. Concordance 4-5 provides more abstract depictions of the concept of surveillance than were found for *forms of surveillance*, but reflects the complexity indicated there. The scholars in the S&S Corpus paraphrase *surveillance* as “a general social phenomenon” (line 13). The variety of surveillance forms is acknowledged in the lines that depict it as “a broad category” or “concept” (lines 1, 2). Even the subcategory of *CCTV surveillance* is described as “a diverse phenomenon” (line 10).

1 e studies acknowledges that surveillance is a broad category that enta 2010-08-1-03
 2 ns to control territory. As surveillance is a broad concept used to de 2005-03-2_3-02
 3 ng we should expect to see. Surveillance is a central organizing princ 2011-08-4-09
 4 t al. 2008). The science of surveillance is a complex equation of huma 2011-08-4-15
 5 priate means of resistance. Surveillance is a condition of modernity, 2011-08-4-07
 6 n, the glassy stare of CCTV surveillance is a constant of urban life. 2004-02-2_3-19
 7 arm in an era in which some surveillance is a constant. To the Surveil 2015-13-1-07
 8 arxist' in highlighting how surveillance is a critical feature of mode 2005-03-1-06
 9 g people. It is argued that surveillance is a crucial aspect of care a 2010-07-3_4-07
 10 ntly been pointed out, CCTV surveillance is a diverse phenomenon (see 2004-02-2_3-07
 11 , 1995). Moreover, although surveillance is a dominant theme in Wester 2008-05-1-01
 12 ecting privacy. And yet, if surveillance is a form of power, it is sur 2011-08-4-07
 13 Surveillance Studies - i.e. surveillance is a general social phenomeno 2013-11-1_2-10
 14 reinforcing the point that surveillance is a global problem that requ 2011-08-4-12
 15 From this perspective then, surveillance is a means of social control 2012-09-4-03
 16 of domination. In practice, surveillance is a mode of governance, one 2010-07-3_4-11
 17 l justice' (Lyon 2003a: 1). Surveillance is a powerful means by which 2012-10-3_4-02
 18 d. Security can be secured. Surveillance is a powerful technology for 2004-02-2_3-05
 19 s to posit that global mass surveillance is a technology of power depl 2015-13-2-03
 20 what of the surveyed life. Surveillance is a technology that operates 2008-05-3-07
 21 irtual space (Taylor 2009). Surveillance is a threat to the integrity 2012-09-4-06
 22 f the following conclusion. Surveillance is a tool, one that sometimes 2003-01-2-02
Concordance 4-5: Examples of *surveillance* is a pointing to the diversity of the concept (20 out of 75 total occurrences of *surveillance* is a)

Surveillance is shown to play an important role in society: it is considered a “condition of modernity” (line 5) and “a central organizing principle of our times” (extended context of line 3). More dramatically, *surveillance* is considered a “global problem” (line 14) and a “threat” (line 21). The concordance further links *surveillance* to “power” and “powerful” measures (lines 12, 17, 18) and describes it as a “mode of governance” (line 16). From a more technical perspective, *surveillance* is seen as a “technology” (lines 18–20), “tool” (line 22), or “a complex equation of human machine interaction” (line 4). Beyond its general role in society, *surveillance* is regarded as an important concern in particular domains: it is described as “a critical feature of modern education policy and schools” (line 8), “a crucial aspect of care” (line 9) and “a threat to the integrity of businesses” (line 21).

1 r...overseer remind us that **surveillance is not** a mere glance exchanged 2005-03-1-07
 2 oreover, in its operation **surveillance is not** just a general 'staring' 2003-01-2-06
 3 iors" (Giddens 1984, 16). **Surveillance is not** just a matter of the gaz 2009-06-3-03
 4 best captured by privacy. **Surveillance is not** just a problem because i 2011-08-4-08
 5 propriations suggest that **surveillance is not** just a steady growing se 2005-03-2_3-07
 6 y. This is why empowering **surveillance is not** just possible, but desir 2010-08-2-03
 7 es and Accenture profits. **Surveillance is not** just technological 'bell 2011-09-1_2-04
 8 rmation gleaned from body **surveillance is not** merely a 'data image', a 2009-06-4-04
 9 g note that "the power of **surveillance is not** merely that it is exerci 2003-01-3-06
 10 agement norms, electronic **surveillance is not** necessarily linked with 2003-01-2-02
 11 e and negative aspects of **surveillance is not** only a conceptually chal 2014-12-2-03
 12 lture. In such a culture, **surveillance is not** only a form of entertain 2015-13-2-02
 13 ws how, in this instance, **surveillance is not** only a matter of control 2009-07-1-05
 14 and workplaces. However, **surveillance is not** only a top-down phenome 2014-12-3-13
 15 ecedented because the new **surveillance is not** only about the internet 2012-10-3_4-11
 16 s. Our suggestion is that **surveillance is not** only an unfortunate "sid 2010-08-2-09
 17 rts to clarify how global **surveillance is not** only confined to intelli 2015-13-2-03
 18 e nature of the exchange, **surveillance is not** only embedded into the s 2010-08-2-04
 19 ates that the practice of **surveillance is not** only reflective of socia 2011-08-4-02
 20 have worked to show that **surveillance is not** only that, but it is als 2015-13-1-06
 21 n's (2007) arguments that **surveillance is not** simply a top-down system 2013-11-1_2-04
 22 . Kernels of Surveillance **Surveillance is not** simply about the power t 2009-07-1-03
 23 ct. In the later instance **surveillance is not** simply exercised over vi 2010-08-2-02
 24 tization of surveillance. **Surveillance is not** simply pervasive, in a s 2009-07-1-03
 25 ersation about privacy or **surveillance is not** simply power but agency. 2011-08-4-10
 26 to intervention. That is, **surveillance is not** simply the act of watchi 2010-08-2-08

Concordance 4-6: *Surveillance is not* + diminishing adjective/adverb + concept (26 out of 79 total occurrences of *surveillance is not*)

Concordance 4-6 shows examples of the negated cluster. The selected lines all contain the pattern *surveillance + is not + adverb/adjective with diminishing function + a concept*. Accordingly, the cluster *surveillance is not* commonly suggests that the mentioned concept is part of the meaning, but does not cover everything. The pattern tends to be followed by an explanation of what components the initial concept is missing. In Example (2), this additional component is the “purpose” for the monitoring and the “judgement” that the subject experiences. Example (3) argues that technology is not enough: in addition, bureaucratic control is crucial. Both examples suggest that the concept of surveillance is more complex than it may initially seem.

- (2) Moreover, in its operation **surveillance is not** just a general ‘staring’ at the world; it is always with a purpose, i.e. to make some judgement about the one being monitored.
[line 2; 2003-01-2-06; (Introna, 2003)]
- (3) **Surveillance is not** just technological ‘bells and whistles’ but works in tandem with bureaucratic and ‘on the ground’ methods of control.
[line 7; 2011-09-1_2-04; (Maki, 2011)]

These concordance examples of *surveillance* have indicated that it is described as a complex concept that can take many forms; a general social phenomenon that is a widespread and important aspect of society. Due to these characteristics, surveillance is presented as a topic worthy of academic study. The diverse aspects of surveillance also shape the academic discipline: while surveillance studies is often referred to as a *field*, its interdisciplinarity is a key feature. Overall, the *surveillance* clusters in this section have provided a good starting point for examining paraphrases of *surveillance* in order to explore its meaning components.

4.2.2 Explicit definitions of *surveillance*

From the complexity shown in Section 4.2.1, it follows that there is a need for the scholars writing for the S&S journal to state explicitly how they understand surveillance and “operationalise” it for their study. The clusters analysed in the previous section have pointed to paraphrases of particular meaning components of *surveillance*. In this section, I focus on a more specific subset paraphrases that are explicitly marked as definitions of *surveillance*. I argue that definitions of *surveillance* are a good site for Teubert’s (2010, p. 129) concept of ‘meaning negotiation’.

Another study that illustrates the negotiation of meaning is Mahlberg’s (2007c) work on *sustainable development* (SD; see Chapter 2). She cites a definition of SD from an influential report that was quoted on the website of an authoritative body, the UN’s Division for Sustainable Development (Mahlberg, 2007c, p. 198). We may interpret the fact that the UN posted this definition as a signal that the meaning of SD is fixed and unquestioned. However, Mahlberg’s (2007) case study of *Guardian* articles reveals a wide range of meaning components in the concordance lines of SD. Like *surveillance*, SD is a broad concept. Mahlberg (2007c, p. 199) is able to distinguish various meaning aspects of SD in her specialised corpus by grouping

concordance lines according to ‘local textual functions’. One of these groups, “SD means”, features explicit discussions about the meaning and use of SD, including criticism of it being an “irrelevant buzz phrase” (Mahlberg, 2007c, p. 205).

In a similar way to the negotiation of SD, it can be anticipated that contributors to the S&S Corpus discuss the meaning of *surveillance*. Because publishing “innovative” research on surveillance is part of the editorial aims of the S&S journal (see Section 3.2.1), this innovation will be reflected in the discourse. Meaning-making principle (i), as set out in Section 2.7, states that meaning evolves with the discourse and therefore every occurrence of *surveillance* contributes to its meaning (see Teubert, 2010, p. 180). As methodologies and theoretical frameworks develop, social contexts change and new surveillance technologies are introduced, conceptualisations of surveillance can be expected to shift accordingly in the S&S Corpus. Therefore, negotiations about the definition of *surveillance* are likely to form an integral part of this academic discourse.

Definitions of *surveillance* have to facilitate mutual understanding among surveillance studies scholars who are facing at least two challenges: the field’s inherent interdisciplinarity (see Concordance 4-3, Section 4.2.1) and continuous technological developments that may alter the techniques and characteristics of surveillance. Example (4) illustrates a case of disagreement in relation to the definition of *surveillance*. This early S&S article argues that dictionary definitions do not keep up with developments in surveillance studies. The article goes on to argue that the concept is more complex than presented by the dictionary, particularly with regard to who is watching, who is being watched and where the surveillance is taking place. As dictionary definitions aim to generalise across all uses of a word (see Section 2.3; Teubert, 2007b, p. 68), it is not surprising that they are not sufficiently nuanced for the expert community of the S&S Corpus.

- (4) One indicator of rapid change is the failure of dictionary definitions to capture current understandings of surveillance. For example in the Concise Oxford Dictionary surveillance is defined as “close observation, especially of a suspected person”. Yet today many of the new surveillance technologies are not “especially” applied to “a suspected person”.

[2002-01-1-02; (Marx, 2002)]

While the act of defining key concepts serves an important role in academic discourse, relatively little recent linguistic research seems to have focused on the function of academic definitions (cf. Section 2.5.1). An exception is Triki’s (2019) study of definitions in research articles. She argues that definitions do not only have interpersonal and textual functions in academic writing, but can also contribute to meaning-making. Similarly, Pearson (1998) has shown that definitions have a performative function, but that their function depends on a variety of factors including the text type and audience as well as the status of the author. Unlike dictionaries, definitions in the S&S Corpus have a more specialised target audience of experts and their peers, though some will have a higher authority in the field than others.

The conceptual approach taken by Taylor (2008, p. 183) to study how “the specific term *corpus linguistics* is used and defined in practice” provides another reference point for my analysis. Compared to the above-mentioned studies of definitions, Taylor (2008) is less concerned with identifying general formulae for definitions. Instead, her aim is to apply corpus linguistic methods to examine how corpus linguists describe their own academic discipline (see Section 2.5.1). Likewise, my aim in this section is meaning-oriented.

In order to extract definitions, I used *WordSmith Tools 7* to generate a concordance of *surveillance* with *define/defines/defined/defining/definition* within a span of five words (see Appendix C, Concordance 1). The plural form *definitions* was not included, because it is unlikely to refer to specific instances. This form only co-occurs with *surveillance* 13 times

(shown in Appendix C, Concordance 4); some of these are irrelevant or overlapping instances, while others emphasise the variety of definitions in the field or summarise main trends.

The initial query generated 160 concordance lines. I excluded 36 lines from the analysis, which either overlap with other concordance lines or are irrelevant. Example (5) illustrates an “irrelevant” instance in which *defined* co-occurs with *surveillance* in a span of five words, but refers to another noun (*population*). All excluded concordance lines are listed in Appendix C (Concordances 2 and 3).

- (5) It is argued that DNA profiling and databasing enable the construction of a ‘closed circuit’ of **surveillance** of a *defined* population.
[line 12, Concordance 7, 2004-02-1-01; (Robin Williams & Johnson, 2004)]²

This methodology for identifying instances of *surveillance* definitions is not exhaustive. It is quite possible for the speech act of defining to take other shapes than the verb *DEFINE*, for example through forms of the copula *be*, the most common syntactic signal for definitions in Flowerdew’s (1992, p. 212) corpus of science lecture definitions, and verbs like *mean* or *refer*. However, I argue that the *DEFINE* markers highlight paraphrases as particularly important contributions.

The following subsections discuss the three main themes that emerged from the qualitative analysis of the concordance lines: broad definitions (Section 4.2.2.1), specialised definitions (4.2.2.2) and the relation of surveillance to definitions of other concepts including surveillance studies (4.2.2.3). This last section discusses concordance lines that do not directly define the term *surveillance* itself, but still contribute to the wider picture of characterising the concept of surveillance.

² Throughout this thesis, examples of co-occurrence pairs show the node in bold and the collocates in asterisks.

4.2.2.1 Broad definitions of *surveillance*

Definitions are likely to reflect a subjective view of surveillance, which not all members of the surveillance studies discourse community may agree with. Such sites of disagreement are important for the negotiation of the meaning of *surveillance*. One dimension along which disagreement may arise is that of generality, because definitions have to strike a balance between covering all possible cases and becoming too vague. This subsection starts with the discussion of commonly cited definitions and moves on to aspects of debate concerning broad definitions. It ends with a discussion of concordance lines that problematise the notion of a broad definition of *surveillance* altogether.

48	al surveillance Lyon defines	surveillance as "any collection and proce	2009-06-3-03
64	led through data collection.	Surveillance has been defined as: "any co	2010-08-2-03
69	the most cited definition of	surveillance proposed by Lyon (2001) suff	2011-08-3-14
91	f Lyon's definition reveals,	surveillance has the aim of influencing o	2012-09-4-03
92	David Lyon (2001, 2) defines	surveillance as "the collection and proce	2012-09-4-03
93	to Lyon's definition, social	surveillance certainly involves "the focu	2012-09-4-04
94	Overview, David Lyon defines	surveillance as "the focused, systematic	2012-09-4-04
95	o surveillance. Lyon defines	surveillance as "any collection and proce	2012-09-4-06
108	defines as 'social sorting,'	surveillance is never neutral, since it e	2012-10-3_4-09
109	ty and Ericson (2000, 2006),	surveillance has been defined as the '...	2012-10-3_4-09
114	n's (2007) definition: while	surveillance 'is the focused, systematic	2013-11-1_2-04
118	idely accepted definition of	surveillance proposed by Lyon (2007), I d	2013-11-1_2-05
125	ing David Lyon's definition,	surveillance takes the form of routine at	2013-11-1_2-13
129	aracteristic of contemporary	surveillance society (Lyon 2010), or even	2013-11-3-04
132	ple, Lyon (2007: 14) defines	surveillance as 'the focused, systematic	2014-11-4-04
135	, it has been suggested that	surveillance , defined as 'watching over'	2014-11-4-12
140	humans As mentioned before,	surveillance is commonly defined as an 'e	2014-12-2-07
141	in current Big Data trends.	Surveillance is commonly defined as the '	2014-12-2-07

Concordance 4-7: All 18 instances of definitions referring to Lyon's work

Pearson (1998, p. 112) argues that "the audience will determine whether or not the definition is acceptable". This point relates to the broader argument that "[w]e can measure the success of an utterance by the traces it leaves in subsequent utterances" (Teubert, 2010, p. 115). These arguments on intertextuality and meaning in discourse suggest that widely cited definitions position the original text (or author) as an authority in the field. Indeed, many instances in the DEFINE concordance represent an intertextual reference.

Among the academic sources, definitions by David Lyon provide a striking example. The 18 concordance lines (from 14 different articles) citing his work are reproduced in Concordance 4-7. Example (6a) shows a commonly cited definition by Lyon. A comparison of this citation to Lyon's (2001) original text in (6b) supports Pearson's (1998, p. 112) argument that not all definitions are necessarily preceded by a "performative utterance" (e.g. *I define, I declare*). One possible reason for leaving out a performative utterance is the established authority of the writer in the field (Pearson, 1998, p. 112), which appears to be the case for Lyon. Notably, while Lyon's original quote does not include a performative utterance, the rhetorical question "What is surveillance?" similarly prompts peers to understand the following sentence as a definition, leading to the quote being introduced by clauses like *Lyon defines surveillance as* in Example (6a).

(6)

- a) Lyon **defines** **surveillance** as "any collection and processing of personal data, whether identifiable or not, for the purposes of influencing or managing those whose data have been garnered" (2001, 2).
[line 48, Concordance 1, 2009-06-3-03; (Martin, Brakel, & Bernhard, 2009)]
- b) What is surveillance? In this context, it is any collection and processing of personal data, whether identifiable or not, for the purposes of influencing or managing those whose data have been garnered. (Lyon, 2001, p. 2)

(7) [...] Lyon (2007: 14) **defines** **surveillance** as 'the focused, systematic and routine attention to personal details for purposes of influence, management, protection or direction'.

[line 132, Concordance 4-7; 2014-11-4-04; (Whelan, 2014)]

Example (7) cites a later definition by Lyon from 2007, that is referred to several times in Concordance 4-7 (lines 93, 94, 109, 114, 118, 140, 141). Although the overall meanings in Examples (6) and (7) appear to be similar, these definitions differ in the subject of the surveillance activity: *personal data* vs. *personal details*. In the description of definitions, the

term to be defined (in this case *surveillance*) is known as the ‘definiendum’ and the class this term belongs to is the ‘genus’ (Flowerdew, 1991, p. 255). Following this terminology, the genus of *surveillance* also differs between Examples (6) and (7): “any collection and processing” vs. “the focused, systematic and routine attention”.

Another genus appears in line 135 (Concordance 4-7), where the quoted definition from Lyon reads only two words: “watching over”. This example shows that references to previous texts, like Lyon’s work, can be rather selective indicating how intertextuality is evaluative: by “importing” a formulation from another text, the text engages with the attitude from that source (Teubert, 2005a, p. 187).

In the S&S Corpus, the choice of genus appears to be important, at least in general definitions. To some extent, this choice relates to a debate about the visual nature of surveillance. For instance, the article that critiques a general dictionary definition, quoted in Example (4), also criticises the visual genus *observation*, as Example (8) illustrates.

- (8) The dated nature of the definition is further illustrated in its seeming restriction to visual means as implied in “observation”. [...] A better *definition* of the new **surveillance** is the use of technical means to extract or create personal data.
[line 1, Appendix C, Concordance 1, 2002-01-1-02; (Marx, 2002)]

Although this critical view of restricting the concept of surveillance to the visual is already present in the first S&S volume, similar criticism still appears in Volume 13. In an article (2015-13-1-09), which is both an academic reference for the present study (see Section 4.4) and part of the corpus data, Mehrabov (2015) points out that surveillance techniques like voice recognition do not depend on the visual. The definition of *surveillance* in an article from Issue 13(2) still uses the visual genus *a set of watching and data collection practices* (extended context of line 158, Appendix C, Concordance 1; 2015-13-2-09). However, the concordance

lines suggest that the more visual genera like *observation* and *watch* tend to occur in quotes from dictionaries (e.g. Appendix C, Concordance 1; lines 4, 5, 112, 148). The genus *watchfulness* is an exception that is tied to a specific definition of surveillance in a medical context (see Section 4.2.2.2).

The scope of the surveilled subject is another disputable element of surveillance definitions. Whereas Lyon's definitions in Examples (6) and (7) have a clear focus on people, Example (9) illustrates a definition that is deliberately inclusive in terms of the subject being surveilled. It explicitly covers human and non-human surveillance.

- (9) The essay draws on a more radical and explicitly inclusive *definition* of **surveillance** as encompassing “all forms of monitoring and control of human and nonhuman subjects, from individual people and things to groups, ecosystems, and planetary processes” [...] [line 104, Appendix C, Concordance 1, 2012-10-2-02; (Braverman, 2012)]

This extension of surveillance to non-human subjects is contested: Example (10) recommends a focus on human subjects. This instance questions the scope of a surveillance definition in relation to the aims of the discipline. Other concordance lines also imply or acknowledge the complexity involved in defining the concept of surveillance.

- (10) Indeed, investigating the monitoring of a machine in a factory may seem interesting, but is this important for surveillance studies? Does this even fit our *definition* of what **surveillance** is? In my opinion, for now, our starting point must be human. [line 10, Appendix C, Concordance 1, 2003-01-4-04; (Adey, 2003)]

Concordance 4-8 lists examples expressing the difficulty of finding or agreeing on a general definition for surveillance. Some of these explicitly voice concerns over any general, catch-it-all definitions or conceptualisations. In several lines the decisive context is outside the printed concordance characters. For example, the adjectives cut off before the phrase *to define* in lines

53, 54 and 130 are all negative: *unproductive*, *difficult*, and *unable*. Similarly, the context of line 149 refers to the ideal of a general surveillance definition with the sceptical statement “should that indeed be possible”. The extended context of line 55 criticises that definitions of *surveillance* tend to be merged with definitions of “political and social problems”, leading to “conceptual confusion” and argues that clear criteria are needed for defining surveillance. On the other hand, instances from article 2015-13-1-03 further warn of the danger of biasing research with *a priori* definitions.

22	eed upon definition of what surveillance is. Marx's (2005, 2001) iter	2005-03-2_3-03
53	ve to define 'the nature' of surveillance in general terms. Instead, i	2009-07-1-05
54	icult to define. Apparently, surveillance is at work in every corner o	2009-07-1-05
55	a to define surveillance qua surveillance , conceptual confusion will c	2009-07-1-07
67	in particular, how to define surveillance . In this article, we do not	2010-08-2-09
130	able to define the powers of surveillance or, indeed, to devise a mean	2013-11-3-05
149	o secure any definition for ' surveillance ' of course exceeds the scope	2015-13-1-03
151	h fields or sites defined as surveillance in an a priori way (via the	2015-13-1-03
152	hips in settings defined as ' surveillance ' in an a priori fashion. The	2015-13-1-03
153	hips in settings defined as ' surveillance ' in an a priori fashion. As	2015-13-1-03
154	es that does not investigate surveillance defined as a set of 'things'	2015-13-1-03

Concordance 4-8: Instances problematising broad definitions of surveillance

Concordance 4-8 only represents six articles, so criticism of broad definitions is not dominant. Nevertheless, the instances demonstrate that this view has left some “traces” in the discourse (see Teubert, 2010, p. 115). As these examples are mainly found in the latter half of the corpus, this view might be gaining more support.

4.2.2.2 Specialised definitions of surveillance

Some articles employ general definitions as templates for defining specific research contexts or types of surveillance. Example (11) shows how Lyon’s definition is adapted to describe *disability surveillance*. It illustrates an explicitly intertextual relation in the discourse, as a writer uses a previous text to strengthen their argument.

- (11) Building from a widely accepted definition of surveillance proposed by Lyon (2007), I *define* disability **surveillance** as the practice of collecting, documenting, monitoring and classifying personal data that pertains to the embodied characteristics and attributes of impairment.

[line 118, Concordance 4-7, 2013-11-1_2-05; (Saltes, 2013)]

In other cases, the relationship to a general definition of surveillance is minimal in favour of characterising the specialised context. As a case in point, Concordance 4-9 shows definitions of *surveillance* that co-occur with *disease*, *(public) health* or *medicine*. These medical definitions are specialised both in their assertions and their dispersion in the corpus so that they almost appear to form their own sub-discipline.

32	been a defining attribute of surveillance <i>medicine</i> ; now such self exam	2009-06-2-01
33	definition of <i>public health</i> surveillance and its continued relevance	2009-06-2-01
34	two recent documents define surveillance <i>in a public health context</i> i	2009-06-2-02
35	rkelman define <i>public health</i> surveillance , these contributors, whoever	2009-06-2-02
37	following definition: <i>Health</i> Surveillance may be defined as the tracki	2009-06-2-02
38	[...] Langmuir defined <i>disease</i> surveillance as 'the continued watchfulne	2009-06-2-02
41	down the basic definition of surveillance <i>in a public health context</i> .	2009-06-2-02
43	overnance. <i>In public health</i> , surveillance , as defined by Alexander Lan	2009-06-2-03
45	of what has been defined as ' surveillance <i>medicine</i> ': This new surveill	2009-06-2-07
139	ces, including <i>public health</i> surveillance , which has been defined as "	2014-12-2-04
143	l shift in the definition of surveillance <i>in a public health context</i> t	2014-12-2-08

Concordance 4-9: Examples of specific definitions of *surveillance* in the context of (public) health

- (12) In public health, **surveillance**, as *defined* by Alexander Langmuir, means 'the continued watchfulness over the distribution and trends of incidence through the systematic collection, consolidation and evaluation of morbidity and mortality reports and other relevant data' for purposes of prevention of disease or injury' (Langmuir 1992).³

[line 43, Concordance 4-9, 2009-06-2-03; (Bauer & Olsén, 2009)]

There is a thematic difference between general definitions of *surveillance* in 4.2.1 and those in a medical context. Example (12) provides the extended context of line 43 in Concordance 4-9, illustrating the restriction to medical data ("morbidity and mortality reports") and purposes

³ The source article contains two closing quotes as reproduced here.

(“prevention of disease or injury”). Moreover, the sentence structure introducing the definition roots it “[i]n public health” rather than in surveillance studies. Concordance lines 38, 41 and 143 from two other articles also quote Langmuir’s definitions (from the 1960s or 1990s). The extended context of line 37 further shows a lasting impact of Langmuir’s work by referring to a more recent definition that builds on it. This definition is also quoted in Volume 12 (line 139).

All examples in Concordance 4-9 originate from the special issues 6(2) and 12(2) entitled *Health, Medicine and Surveillance* and *Big Data Surveillance*, respectively. Concordance lines 34–41 originate from one article (2009-06-2-01), which explicitly reviews the history of terminology in this area. This restricted dispersion suggests that public health surveillance may form an interdisciplinary sub-field of surveillance studies with its own intertextual relations. Public health surveillance was chosen as a case study for specialised definitions. While some other specialised definitions are mentioned in the context of special issue themes in Section 4.4, others remain for future research to investigate.

4.2.2.3 *Surveillance* and definitions of other concepts

This subsection considers instances from the DEFINE concordance where not surveillance, but other concepts constitute the definiendum. Two groups of examples are discussed: surveillance defining other concepts and definitions of *surveillance studies*.

16	e disciplinary society, such	surveillance captures and defines the sub	2004-02-2_3-19
44	eal time, but via prosthetic	surveillance regulates and defines bodies	2009-06-2-07
46	endency towards a prosthetic	surveillance , which regulates and defines	2009-06-2-07
50	echnologies and practices of	surveillance helped define the national a	2009-07-1-01
63	ation defined by a matrix of	surveillance that includes border patrols	2010-08-2-02
71	y also demonstrates how this	surveillance uses race to define the bord	2011-08-4-02
72	gue that practices of direct	surveillance define the border between th	2011-08-4-02
73	by mediated visibilities and	surveillance , and it is not necessary to	2011-08-4-03
83	is best defined in terms of	surveillance , the social and individual v	2011-08-4-08
84	nnotes the modern landscape.	Surveillance as a definition of, or frame	2011-08-4-08
85	ould be defined in terms of	surveillance because it more accurately f	2011-08-4-12
110	surveillance in daily life,	surveillance stands out as a defining fea	2012-10-3_4-09
122	2011). 2) Normative: liminal	surveillance defines what can be seen as	2013-11-1_2-08
146	participatory nature of the	surveillance was a defining aspect of the	2014-12-3-03

Concordance 4-10: All instances of *surveillance* as a definition for other issues

Concordance 4-10 shows the 14 concordance lines that emphasise the role of surveillance in shaping various social contexts. Here, surveillance is depicted as a framework for affecting and even defining other social concepts, such as “the border between the community and ‘stranger’ Others” in line 72. Surveillance is also described as a “defining aspect” of small- and large-scale cultures: “the ‘safe space’ that was this World of Warcraft guild” (context of line 146) and “a modern global culture” in the extended context of line 73.

31	the definition and scope of Surveillance Studies , looks at surveillan	2009-06-1-08
58	efine the next generation of Surveillance Studies , and I offer encoura	2009-07-1-08
59	enting something defined as ' Surveillance Studies ' - matters to the fu	2009-07-1-08
87	e spend less effort defining surveillance studies , as this literature	2012-09-3-05
123	of policing. By definition, Surveillance Studies is concerned with al	2013-11-1_2-10

Concordance 4-11: All concordance defining or negotiating the definition of *surveillance studies*

The second group of examples directly defines the disciplinary nature of *surveillance studies*. In view of the editorial scope of the S&S journal, described in Chapter 3, this question is important for all publications in the journal. Various examples discussed throughout Section 4.2 have highlighted disciplinary aspects, such as Example (10) questioning the significance of non-human surveillance for the discipline. Concordance 4-11 shows instances from the DEFINE concordance that define *surveillance studies* rather than *surveillance*. These notions are connected, because the definition of *surveillance* inherently links to the objectives of the scholarly field and what it investigates (see Section 4.2.1).

The article to which lines 58 and 59 belong is a sequel to a critical book review, responding to the book authors’ negative reaction. As such, it clarifies the rationale behind the original critique; Example (13) illustrates that the positioning of the discipline is of major concern.

(13)the framing - what one chooses to present as representing something **defined** as ‘**Surveillance** Studies’ - matters to the future of the field [...] if one sets one’s boundaries by a discipline (in this case, Sociology), one cannot hope to encompass a field that is interdisciplinary or transdisciplinary.

[line 59, Concordance 4-11, 2009-07-1-08; (Murakami Wood, 2009a)]

(14)By **definition**, **Surveillance** Studies is concerned with all dimensions of surveillance and therefore encompasses a great variety of questions (modalities, functions, effects, driving forces, limits and so forth) in a cross-disciplinary perspective.

[line 123, Concordance 4-11, 2013-11-1_2-10; (Germain, Dumoulin, & Douillet, 2013)]

Similarly, the context of line 123 represents surveillance studies research as crossing the boundaries of particular disciplines. Example (14) shows that this perspective on the field is closely tied to the understanding of the concept of surveillance, namely that it has many “dimensions” that can be investigated through a wide range of research questions (see Section 4.2.1). The two remaining lines contain metalanguage that does not accompany actual definitions. Line 31 is part of a book review, stating that the book defines surveillance, and line 87 actually suggests that the readers of the articles do not require an extensive definition of the field of surveillance studies.

This section has demonstrated that defining *surveillance* is a multifaceted activity in the S&S Corpus. While a set of definitions is quoted repeatedly, the negotiation about definitions is still at the heart of the discipline. Some disagreement emerges regarding components of broad definitions, and more specific definitions are put forward. Accordingly, the S&S journal does not present one single view of surveillance, but provides a platform for negotiation. Contested aspects include the scope of surveillance (e.g. whether it extends to monitoring non-human subjects), its restriction to visual elements and the risk of a priori definitions skewing research findings. The section has further shown that in this corpus surveillance plays a major role in defining other social concepts and the field of surveillance studies itself.

4.3 Key keywords in the S&S Corpus

The clusters and definitions analysed in the previous section provide an approach to examining meaning that describes meaning through lexicogrammatical context. That perspective is complemented in this section by looking at salient words in the corpus (key keywords). *Surveillance* is unsurprisingly one of them, but the keyword analysis also reveals other important words. Section 4.3.1 outlines the procedure for identifying key keywords and Section 4.3.2 discusses the meaning groups that they form.

4.3.1 Identifying key keywords

Key keywords (KKWs) are one level above “simple” keywords: as explained in Section 2.4.5, KKW analysis looks for shared keywords across a given number of subcorpora. In this study, I work with keywords for each of the volume subcorpora compared to the (original) British National Corpus (see Section 3.3.2).⁴ For a word to be included in the KKW list, it has to be identified as a keyword in all 13 volumes of the journal. This procedure generates a list of words that are salient across the full corpus.

WordSmith Tools (Version 7.0.0.45) was used for the KKW analysis. Figure 4-1 shows the stages of the KKW analysis. First, the original UTF-8 text files were converted to a *WordSmith Tools*-preferred encoding format using the *WordSmith Tools 7* internal Text Converter tool.⁵ In step (ii) I used the *WordList* tool to generate one wordlist for each volume. I adjusted the “language settings” for wordlists to consider hyphens, apostrophes and periods as valid characters within words (“Hyphens separate words = false” and “characters within

⁴ For the comparison I used the wordlist of the “BNC World Edition” (whole corpus), available from the *WordSmith Tools* Website (Scott, n.d.).

⁵ The following text converter options were activated: “whole files; into Unicode; .TXT file-extensions; curly quotes etc.”.

word = ‘.’). These settings count word forms such as *sub-heading*, *i.e.* and *don’t* as single words.

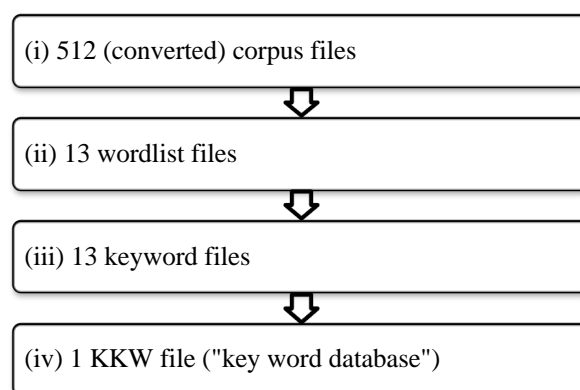


Figure 4-1: Diagram representing procedure for identifying KKWs in *WordSmith Tools*

For identifying the keywords in step (iii), a low *p*-value cut-off of 0.000001 and a minimum frequency threshold of three were used. The minimum log ratio value was set to two, ensuring that all keywords for the S&S Corpus would be at least four times as frequent as in the BNC. The “maximum wanted” value was set at 16,000; high enough to accommodate the keywords for each volume without introducing another arbitrary threshold. For determining KKWs, the “minimum texts for database” was set to 13. “Text” here refers to one volume rather than an individual article, so this setting guaranteed that each KKW had to be “key” in every volume and would have a high dispersion level throughout the corpus.⁶ These settings yield 69 KKWs (see Table 4-3, row 1) in step (iv). Without the log ratio threshold, 117 items are found; Table 4-3 also lists the 48 excluded items.

⁶ Although *WordSmith Tools* offers two other dispersion settings, these were deactivated, because their effect was not considered transparent enough. These functions are “min. % of texts”, which disregards any keyword candidate that does not appear in more than a set percentage of all files, and “min. KWs per texts”, which ignores any keyword file with fewer than the given number of keywords when determining the KKWs.

Table 4-3: KKWs in the S&S Corpus (key in all 13 volumes, compared with the BNC)

All 69 KKWs (log ratio ≥ 2)	abstract, agencies, al, argues, automated, behavior, camera, cameras, cctv, citizens, concerns, contemporary, contexts, data, databases, deleuze, digital, disciplinary, discourses, empirical, enforcement, ericson, et, eu, everyday, focus, forms, foucauldian, foucault, foucault's, governance, i.e, identification, identities, identity, individuals, institutional, internet, lyon, mechanisms, monitor, monitored, monitoring, networks, normative, online, panoptic, panopticon, perceived, policing, populations, practices, privacy, profiling, regarding, regulation, security, social, space, spaces, surveillance, technological, technologies, terrorism, theoretical, tracking, visibility, visible, websites
48 items excluded (log ratio <2)	access, actions, activities, also, analysis, and, are, article, as, context, control, crime, critical, cultural, example, in, increasingly, individual, information, interaction, issues, law, media, of, often, or, police, potential, power, private, process, processes, protection, public, research, studies, such, systems, techniques, technology, these, this, through, thus, tools, understanding, visual, within

As expected, the KKW list contains the word *surveillance*. In addition, there are several related words such as *monitoring* and *tracking* and words that relate to surveillance measures and *technologies* as well as theoretical concepts (e.g. *panopticon*), among others. The following section explains how I grouped the KKWs further.

The nature of the BNC as a reference corpus has a bearing on this list. For example, the noun *internet* has likely become more frequent since the compilation of the BNC.⁷ Additionally, the noun *behavior* is probably found to be “key key” largely because of its low frequency in the British English reference corpus. Individual cases like these do not have a strong effect on the KKW categories that are described in the next section.

4.3.2 Grouping KKWs

The aim of this section is to summarise the KKWs in the S&S Corpus via meaningful groups. On a prima facie basis, some KKWs form groups of shared meanings: e.g. *monitoring* and *tracking* would belong to a “surveillance” group, whereas *contemporary* and *everyday* would

⁷ As the BNC2014 is not available yet, no direct comparison is possible. In the Corpus of Contemporary American English (Davies, 2008), however, *internet* has a frequency of 6.87 per million (pm) in 1990–1994, overlapping with the BNC, and much higher frequencies in the sections overlapping with the S&S Corpus, in fact peaking in the early 2000s: – 123.70pm in 2000–2004; 95.68pm in 2005–2009 and 79.15pm in 2010–2014.

belong to a different group related to time. Other word forms, like *abstract*, are polysemous and therefore harder to categorise. At this stage, however, it is not desirable to conduct an in-depth concordance analysis for all 69 KKWs, as they appear in too many concordance lines for manual analysis (collectively, 99,245 instances). Instead, Section 4.4 reports on a detailed co-occurrence comparison of the most frequent KKWs.

Table 4-4: KKWs categorised by *USAS* tags

<i>USAS</i> tag category	KKWs
A General & Abstract Terms	abstract, automated, enforcement, forms, practices, privacy, theoretical
C Arts & crafts	camera, cameras
E Emotional actions, states & processes	concerns
G Govt. & the public Domain	citizens, disciplinary, governance, policing, regulation, security, terrorism
I Money & commerce	agencies
M Movement, location, travel & transport	panopticon
N Numbers & measurement	space, spaces
O Substances, materials, objects & equipment	contexts, mechanisms
Q Linguistic actions, states & processes	argues, cctv, discourses, profiling
S Social actions, states & processes	behavior, identities, identity, individuals, institutional, networks, normative, populations, social
T Time	contemporary, everyday
X Psychological actions, states & processes	data, databases, empirical, focus, identification, monitor, monitored, monitoring, panoptic, perceived, surveillance, tracking, visibility, visible
Y Science & technology	digital, internet, online, technological, technologies, websites
Z Names & grammatical words	al, deleuze, ericson, et, eu, foucauldian, foucault, i.e., lyon, regarding

In order to reduce subjectivity in grouping the KKW *prima facie*, they are automatically tagged with semantic tags using the UCREL *USAS* English online tagger.⁸ As explained in Section 2.4.6, the *USAS* tag system contains 21 main categories denoted by letters of the alphabet, which are then further subdivided. Table 4-4 shows the KKWs grouped by their main *USAS* tags. This automatic categorisation is a useful starting point, because it provides an external perspective

⁸ Available from <http://ucrel.lancs.ac.uk/usas/tagger.html>, accessed August 2016.

on this corpus and shows, for example, that the categories of “government and the public domain” and “social actions, states and processes” are salient. As the *USAS* tagger works most accurately on full text, the annotation of the KKW list is not perfect.

For the grouping to be useful for the present analysis, some manual adjustments are carried out that take the context of the S&S Corpus into account. For example, the *USAS* tagger identifies *abstract* as an adjective, although it is mostly used as a noun denoting an article section in the S&S Corpus (304 out of 424 instances). Whilst *camera* and *cameras* might indeed refer to forms of art in the journal, this is unlikely, because most artistic articles, originally published in Issue 7(2), have been removed in the cleaning stage (see Section 3.2.1). These words are more likely to refer to the technology, and the same is true for *CCTV*, which *USAS* has classified as a “linguistic process”. Similarly, for this text type, the word *agencies* is more likely to be related to the government rather than to commerce. According to *WordSmith Tools*, the most frequent L1 collocates are *enforcement*, *security*, *intelligence* and *government*.

Table 4-5: KKW groups

KKW group	KKWs
1. Theoretical frameworks	concerns, contexts, discourses , enforcement, focus , forms, mechanisms , networks, normative, panoptic, panopticon , practices, privacy, theoretical
2. Government & the public domain	agencies , citizens, disciplinary, eu, governance, institutional , policing, regulation, security, terrorism
3. Time & space	contemporary, everyday, space, spaces
4. Research & academic writing	abstract, al , argues, empirical, et, i.e., regarding
5. Social actions & actors	behavior, identities, identity, individuals, populations, social
6. Monitoring & identification	identification, monitor, monitored, monitoring, perceived, profiling , surveillance, tracking, visibility, visible
7. Technology	automated, camera, cameras, CCTV, data, databases , digital, internet, online, technological, technologies, websites
8. Theorists	deleuze, ericson, foucauldian, foucault, faucault’s⁹, lyon

⁹ This was not included in Table 4-4, because the *USAS* English tagger splits words at apostrophes.

Table 4-5 shows the groups after manual adjustment; unsuitable allocations have been manually re-arranged according to contextual information from the corpus and in consultation with the secondary *USAS* tags. In addition, categories containing only one or two KKWs (categories “C”, “E”, “I”, “M”, “N”, “O” and “T” in Table 4-4) were eliminated. KKWs that have changed categories are shown in bold.

These manually adjusted “groups” give a more transparent overview of KKWs in the S&S Corpus. Group 1, “Theoretical frameworks” is broadly based on USAS category “A” and contains references to theoretical notions, such as “a normative theory of surveillance”, which is discussed in a debate section of Issue 12(1), and *panoptic/panopticon*. The *panopticon* is a concept derived from Bentham’s architectural design of a prison in which all prisoners are visible for the central watch tower. Notably, the prisoners do not know if and when they are being watched. Foucault made popularised the panopticon as a theoretical concept (see e.g. Lyon, 2007). *Privacy* is another theoretical concept that is of particular importance in relation to surveillance – and not uncontested, see Example (15).

(15) Notwithstanding sustained critique, the concept of privacy remains a central organising principle within Surveillance Studies (Bennett 2012)
[2013-11-3-04; (Edmond & Roque, 2013)]

Group 2 is modelled on USAS category “G”. While no words have been removed, *agencies* has been added, based on its collocates. *Institutional* also fits in better with this group. Example (16) illustrates that *institutional* refers not only to governmental institutions. Therefore, the “public domain” of group 2 also involves corporations like the *insurance companies* and *marketers* in (16). No other KKW specifically refers to such “private” institutions, which is somewhat surprising, given the increasing spreading of surveillance measures beyond governmental bodies indicated in Section 2.2. However, these references may be more localised

in particular volumes and special issues of the S&S Corpus, such as Issue 1(2) with the theme “Work” and 8(3) titled “Marketing, Consumption and Surveillance”.

- (16) Actuarial surveillance as a technique of knowledge production and population management is becoming a central organizing principal of modern institutions. It is being adopted in more and more **institutional** settings, from insurance companies to marketers to police agencies.

[2015-13-1-05; (D. J. Phillips, 2015)]

Group 3 combines the USAS categories “N” and “T” for a common group of KKWs related to time and space. This group is important for surveillance discourses, because temporal and locational information is central to virtually all types of surveillance measures (see Section 2.6; Lyon, 2007, p. 16). Group 4 largely relates to the genre of academic writing. It includes arguably typical academic function words (*al, et, i.e., regarding*) from the “grammatical words” portion of USAS category “Z” and the word forms *argues* (formerly category “Q”) and *empirical* (from category “X”), associated with intellectual interpretation/argumentation and empirical data collection in research, respectively. Group 5, based on the USAS “S” category, is concerned with society, including, e.g. general references to groups of people (*individuals, populations*) and social concepts (*behavior, identity, social*). This group indicates that despite the technological aspects often associated with surveillance, people are involved and affected, whether it is the policy makers and security personnel or the citizens who are targets and beneficiaries of surveillance. The concept of the interaction order in surveillant landscapes (Jones, 2017), introduced in Section 2.6.1, focuses on these social relationship brought about by surveillance. I apply this framework in Chapter 5.

Groups 6 and 7 are perhaps most prototypical for the lexis surrounding processes of surveillance and related technologies. “Monitoring and identification” is a more relevant label

for group 6 in this corpus than the original USAS “X” label “Psychological actions, states & process”. This group contains several words that might be considered synonyms for surveillance in particular circumstances (*identification, monitoring, profiling, tracking*) and words concerned with *visibility*, which relates to the meaning of the French word *surveiller*, “watching over” (Albrechtslund, 2008, online). Some definitions mentioned in Section 4.2.2 – especially those from dictionaries – also referred to surveillance as a visual concept.

Group 7, then, presents technological terms that are salient in the S&S Corpus. Camera-related technology is better placed here than in an “art” group. I have included *data* and *databases* here because of their technological senses, although they are, of course, often part of the identification procedures listed in group 6. Finally, group 8 highlights the proper names of important historical and contemporary surveillance theorists: Gilles Deleuze, Richard Ericson, Michel Foucault, and David Lyon, whose influence on surveillance definitions I discussed in Section 4.2.

So, this grouping has achieved a high-level overview of the aboutness of the corpus. Overall, the KKWs relate to types of words that generally become keywords (see Scott & Tribble, 2006, p. 55): content words (from most groups) and proper nouns (mainly group 8) indicating the topics of the corpus and words indicating a certain style (in particular group 4). This section has addressed Research Question 1-2 by identifying words that are consistently salient across the S&S Corpus, KKWs generated in comparison with the BNC. Eight KKW groups have been identified that broadly relate to the description of theoretical frameworks of surveillance, (governmental) institutions and groups in society, the coordinates of time and space, academic discourse, social actors, monitoring, technologies, and notable theorists in surveillance studies.

4.4 Differences between volumes

This section addresses the third research question, “How do the meanings of the consistently salient words differ across the journal volumes?”, by examining differences in co-occurrence patterns across the corpus. The KKWs identified in the previous section form the starting point for this analysis, because they are “salient” words across all volumes of the journal in that they have been found to be key in each volume compared to the BNC. The shared salience means that the KKWs provide a lexical presentation of high-level aboutness across the corpus. There are parallels between KKW groups 1 and 2 “theoretical frameworks”/ “government & the public domain” and Mehrabov’s (2015) theme of “Classic Surveillance”; KKW group 3 “time and space” and his theme of “Mobility and Stasis”; and thirdly KKW group 5 “social actions and actors” and his theme of “Identity-based Surveillance”. Given that his themes have been derived from special issues, they are localised themes than the KKW meaning groups, which are based on words that are salient across the entire corpus. So Mehrabov’s (2015) final theme of “Work Power and Resistance” is not reflected in the KKW groups. This section examines the co-occurrence patterns of the KKWs in relation to Mehrabov’s (2015) themes. In addition to the analysis of definitions in Section 4.2, the co-occurrence patterns of the KKWs therefore provide a different perspective on the discursive representation of surveillance in the S&S Corpus. Rather than examine the corpus as a whole, this section focuses on specific volumes in order to facilitate a more contextual approach to the various themes to which it is linked in the volumes. In line with the meaning-making principle (i), that meaning evolves with the discourse, the assumption is that the representation of surveillance shifts across the corpus. Section 4.4.1 briefly introduces the four main themes Mehrabov (2015) has identified in the S&S journal based on content analysis. Section 4.4.2 discusses the co-occurrence comparison of the KKWs and explains the results for three illustrative volumes.

4.4.1 Mehrabov's (2015) content analysis themes

This section introduces the four broad themes that Mehrabov (2015) finds in his content analysis of S&S special issues. Table 4-6 lists the four themes along with the S&S special issues that Mehrabov (2015) allocates to them (for the allocation with full special issue titles see Appendix D).

Table 4-6: S&S corpus volumes issues arranged by Mehrabov's (2015, pp. 119-122) themes

Mehrabov's (2015) theme	<u>Volume</u>												
	1*	2	3	4	5	6*	7	8	9	10	11	12	13*
1. Classic Surveillance: "Discipline" and "Control"	1(3)	2(2/3)				6(1)							
2. Identity-based Surveillance				4(3) 4(4)	5(3)	6(2) 6(4)	7(3/4)		9(4)				
3. Mobility and Stasis	1(4)				5(2)				9(3)				
4. Work, Power and Resistance	1(2)	2(4)				6(3)		8(2) 8(3)					

Mehrabov (2015, p. 119) describes the first theme – “Classic Surveillance: ‘Discipline and Control’” – as

part of more traditional Surveillance Studies, existing even long before the launch of the journal [...] and based more on the concepts of Panopticon-related discipline societies (Foucault 1991), and/or “societies of control” (Deleuze 1992).

As indicated in Table 4-6, three issues were allocated to this theme during the content analysis: Issue 1(3) on Foucault and panopticism – due to its link to surveillance as discipline and control, Issue 2(2, 3) on the politics of CCTV and Issue 6(1) on revisiting video surveillance. The use of CCTV cameras is related to “Classic Surveillance”, because it was already studied before the journal launched as a widespread method of watching and controlling populations (see Mehrabov, 2015, pp. 119–120).

Seven special issues across Volumes 4, 5, 6, 7 and 9 are allocated to the second theme, “Identity-based Surveillance”, as shown in Table 4-6. Research in this theme is concerned with social groups that tend to become the target populations of surveillance:

[H]ow surveillance is tracking and affecting lives of elder, children, women, unhealthy, homosexual, homeless people, racially different and immigrant populations—in brief, all the “others” of different social, cultural and economic contexts. (Mehrabov, 2015, p. 120)

The third theme, “Mobility and Stasis” is concerned with surveillance in relation to travel and movement, but also the opposite, being stationery. As Mehrabov (2015, p. 121) states, “the means of transport for this mobility [...] are also rapidly turning into the agents of surveillance, tracking and keeping eye on their passengers and settlers, always in the name of elevated sense of safety”. The first issue related to this theme, 1(4), specifically focuses on mobilities. Issue 5(2) refers back to this theme with a particular focus on the relationship between mobility and borders. Mehrabov (2015, p. 120) also allocates Issue 9(3) on urban surveillance to this issue.

Finally, the fourth theme in Mehrabov’s (2015, p. 122) analysis relates to “the troubled relationship of surveillance with the changing and fixed notions of work and labor” and the potential for resistance and empowerment in these areas. Special issues in volumes 1, 2, 6 and 8 are allocated to this theme in relation to workplace surveillance, surveillance as empowerment and resistance to surveillance. Mehrabov (2015, p. 122) includes Issue 8(3) on surveillance in the context of marketing and surveillance here, because

the internet is [...] a space of close surveillance and data gathering, and [...] this user-generated data is more and more transformed into sellable commodities and the work of internet users [...] is now referred to as “exploitation 2.0” (Andrejevic 2009).

There are methodological differences between Mehrabov’s (2015) analysis and the results I report in the following sections. Most importantly, the themes put forward by Mehrabov (2015) result from a qualitative content analysis (the exact procedure of which is not described in

detail), whereas the results from the present study are based on the corpus linguistic methods of co-occurrence comparisons and concordance analysis of KKWs (i.e. words that are salient in the S&S journal). In addition, the studies differ in the unit of analysis: unlike Mehrabov (2015), who allocates individual issues to themes, I work with entire volumes. Finally, Mehrabov's analysis covers only a subset of the S&S Corpus, because he analyses special issues (rather than open issues) of Volumes 1–9.¹⁰

Despite these methodological differences in the analysis of the journal texts, the four themes put forward by Mehrabov (2015) provide a useful reference point for my corpus linguistic analysis. The fact that Mehrabov's (2015) analysis was published in the S&S journal as a review article – his article is also included in the S&S Corpus as file 2015-13-1-09 – further mean that the arguments have been peer-reviewed by the surveillance studies community. The four themes provide a reflection on the discipline from the perspective of a surveillance studies scholar. So, they give an insight into the development and history surveillance studies that I do not have as a corpus linguist.

4.4.2 Comparing co-occurrences of KKWs

Sections 4.2 and 4.3 have analysed the corpus as a whole entity. References to time were made when relevant to the concordance analysis rather than specifically examined in the methodology (see e.g. the citations of Lyon's 2001 and 2007 definitions of *surveillance* in Section 4.2.2). This section provides a more fine-grained analysis of the volume-based subcorpora in order to examine shifts in the discursive representation of surveillance in the journal. It would seem

¹⁰ Issue 10(1) is also mentioned in relation to the CCTV research of theme 1, but not explicitly allocated to it.

unusual for a journal specialising in analysing a particular social concept not to reflect any development in its theoretical conceptualisation in over a decade (in this case, 2002 to 2013).

The notion of ‘discursive shifts’ is here operationalised to mean changes in the co-occurrence profiles of words that are central to the discourse. The KKWs identified in Section 4.3 are considered to have this salient function, as they have been found to be significantly more frequent in each volume of the S&S Corpus compared to the BNC. Rather than compare all 69 KKWs, I focus on the eleven most frequent KKWs – *surveillance* in addition to the ten most frequent other KKWs – in order to facilitate a more detailed discussion of individual examples.

The methodology for this stage of the analysis applies the co-occurrence comparison method from *CorporaCoCo* (unpublished version 0.13), which has been introduced in Section 3.3. Here, the method is used to identify meaning shifts in the use of KKWs across the volumes of the S&S Corpus in order to address RQ 1-3. For the co-occurrence comparison it seems reasonable to treat *surveillance* separately from the remaining KKW nodes. Ranked among the 20 most frequent words in the corpus (as shown in Section 4.2) and with a frequency of over 20,000 occurrences, *surveillance* has a marked status. As Table 4-7 shows, *surveillance* is over three times as frequent than the next most frequent KKW, *social*. Overall, the nodes cover five out of the eight KKW meaning groups (Section 4.3.1, Table 4-3): “theoretical frameworks” (*practices; privacy*), “government & the public domain” (*security*), “social actions and actions” (*individuals*), “monitoring & identification” (*surveillance*) and “technology” (*cameras; CCTV; data; technologies*). So, these KKWs link to Mehrabov’s (2015) themes of “Classic Surveillance” and “Identity-based Surveillance” (see Section 4.4). In addition, their co-occurrences link them to further themes, as the following subsections demonstrate.

Table 4-7: Node sets for the co-occurrence comparison of the S&S Corpus

Node set	Node	KKW rank	Frequency
1	surveillance	1	20,265
	social	2	6,297
	data	3	6,104
	security	4	4,375
	CCTV	5	4,297
2	privacy	6	3,771
	technologies	7	3,219
	practices	8	2,529
	individuals	9	2,357
	cameras	10	2,183
	space	11	2,159

The following subsections discuss the co-occurrence results for three of the journal volumes: 1, 6 and 13. Volumes 1 and 13 represent the first and final volumes of the journal and therefore have the widest temporal gap, which facilitates snapshots of the meaning aspects of *surveillance* from the beginning and end of the corpus. In addition, I have chosen Volume 6, because it marks the relaunch of the journal on a new website and is therefore an editorial milestone. As Volume 6 is located (almost) in the middle between Volumes 1 and 13 on the publication schedule, it also facilitates an additional insight into the development of the corpus. I discuss the results for these volumes in relation to Mehrabov's (2015) themes. This approach of matching corpus data to an existing, non-linguistic framework is similar to McEnery's (2006) methodology of populating the moral panic theory categories with key keywords. The present case is slightly different, as Mehrabov's (2015) themes are observations about existing research rather than a full-blown theory. Nevertheless, the themes usefully guide the analysis of co-occurrence patterns and represent a way of engaging with expert recommendations in analysing a corpus (rather than compiling it; cf. Section 2.4.2.1). Individual result plots for each of the remaining volumes are provided in Appendix E. For reasons of space they are not discussed specifically, but their results are accounted for in Section 4.5 on co-occurrence links across the full corpus.

4.4.2.1 Volume 1 (2002–2003): The launch

Volume 1 contains four issues, the launch issue, 1(1), with an open topic, and three special issues. The results for the co-occurrence comparison of Volume 1 against the full corpus are shown in Figures 4-2 for *surveillance*. Figure 4-3 gives the significantly different collocates for the second set of nodes. Not all nodes in this set show significant differences (e.g. *individuals* is missing from this plot). Various collocates highlighted by the two plots, especially the collocates of *surveillance*, are related to the special issue titles and can therefore also be linked to Mehrabov's (2015) themes. I discuss these themes in turn.

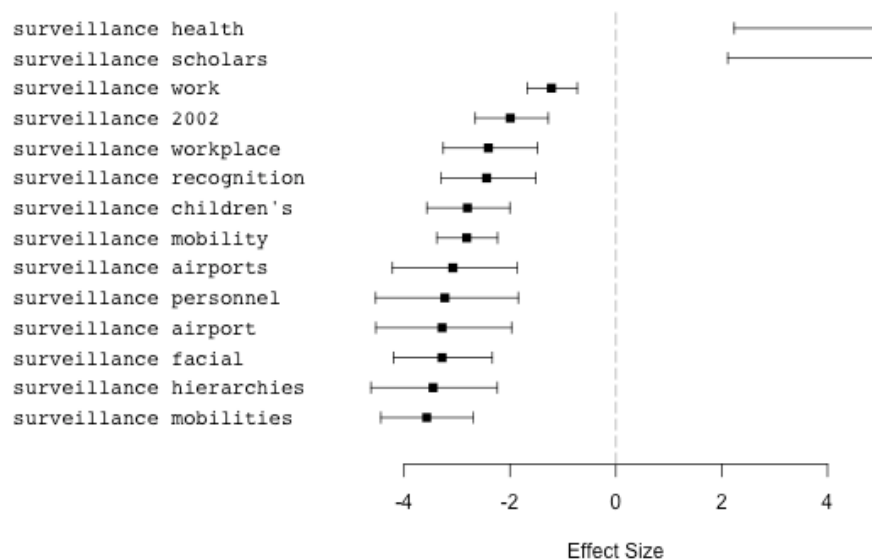


Figure 4-2: Co-occurrences of *surveillance* in Volume 1 vs. the whole S&S Corpus

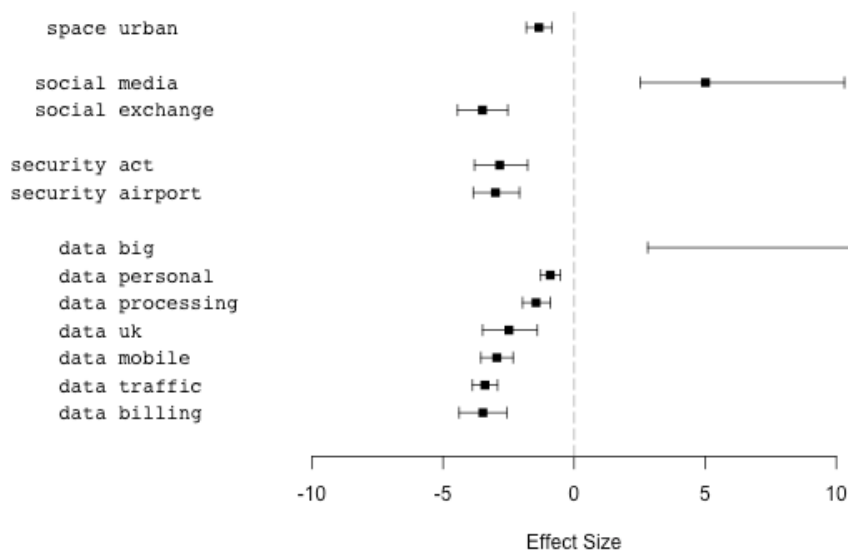


Figure 4-3: Co-occurrences of other KKWs in Volume 1 vs. the whole S&S Corpus

Mehrabov (2015) matches Theme 1, “Classic Surveillance”, with Issue 1(3) “Foucault and Panopticism Revisited”. Out of the three themes to which the issues of Volume 1 have been allocated, I have identified the most *surveillance* collocates for this theme. *Hierarchies* is the second most distinct collocate of *surveillance* in Volume 1 compared to the full corpus (see Figure 4-2). It relates to the traditional view of surveillance as carried out in a top-down manner by the state. The collocate is most frequent in an article which argues that although contemporary theories portray surveillance as carried out by many actors, the situation is “far from a partial leveling of hierarchies” (2003-01-3-10; Hier, 2003). Accordingly, the state may even make use of the other actors in the name of such causes as “safety, security and anti-terrorism” (2003-01-3-10; Hier, 2003).

A second example for this theme is found in the *surveillance* collocates *facial* and *recognition*, which together form the cluster(s) *facial (biometric) recognition (surveillance)*. Like *hierarchies*, these collocates are concentrated in one article. That article discusses this surveillance technique, arguing that it “promises to bring the disciplinary power of panoptic surveillance envisioned by Bentham [...] into the contemporary urban environment” (2003-01-

3-06; Gray, 2003). A final example of the engagement with “Classic Surveillance” is found in the collocate *personnel*, which in the Volume 1 is mostly used as part of the compound *surveillance personnel* (referring e.g. to security guards). In one article in particular, surveillance personnel is discussed as the target of ‘sousveillance’ (acts of counter-surveillance), which therefore works against classic forms of surveillance as discipline and control.

While these collocates discussed for the theme of “Classic Surveillance” mainly occur in individual articles of Issue 1(3), their concordance lines provide lexical support for the theme. The examples further show evidence for ongoing negotiation of what these “classic” concepts mean and how they can be applied to modern contexts – including new technologies like *facial recognition* – or how subversive concepts like *sousveillance* interact with traditional frameworks.

Mehrabov (2015) has not allocated any issue from Volume 1 to the theme of “Identity-based Surveillance”. One co-occurrence pair in Figure 4-2, however, seems to relate to this theme: (*surveillance, children’s*). This pair is not representative of a full issue but just occurs in one article of Issue 1(4), the special issue titled “Surveillance and Mobilities”. The most frequent cluster in which this pair occurs is *surveillance of children’s mobility*. This cluster then takes us to the third theme in Mehrabov’s (2015) overview, “Mobility and Stasis”, to which he has allocated Issue 1(4). The collocate with the biggest difference between Volume 1 and the full corpus is *mobilities*. Its singular form is also significantly more frequent, though the difference is not as large. The plural form actually mainly (20 out of 26 total instances) occurs in the editorial of Issue 1(4), which introduces the general background and the various forms of *mobilities*. The editorial argues that *the surveillance of mobilities* has a profound impact on everyday life and how it is to be understood in surveillance studies, as shown in Example (17).

- (17) The **surveillance** of *mobilities* defies the contextualization of life: the workplace, store and home are no longer separate places in which one is surveilled but instead each becomes a point on the flow of surveillance.

[2003-01-4-01; (Bennett & Regan, 2003)]

The singular form can be used in a similarly general way. Example (18) not only argues that *mobility* is a growing application area of surveillance, but also highlights the connection between mobility and *space*, a KKW in the S&S Corpus. Location is important information for most surveillance systems (see Section 4.3.1; also see Chapter 5).

- (18) Surveillance is increasingly focused upon mobility. Be it in cities, shopping malls or outdoor ‘public’ spaces, surveillance is now able to track and monitor peoples movements.

[2003-01-4-04; (Adey, 2003)]

The cluster *surveillance of children’s mobility* points to a case study of surveillance and mobility. Example (19) observes that surveillance relates to “both care and control” and argues that mobility surveillance has the potential to affect the “perception of space and place”.

- (19) [...] surveillance has two faces and is a matter of both care and control [...] Seen in this light, the **surveillance** of *children’s* *mobility* is not that clear-cut. Have dangers in neighbourhoods and cities increased so much that children need to be monitored in order to care for them properly, or is the monitoring of children’s mobility done on behalf of parental perception with negative consequences to children’s perception of space and place as a result?

[2003-01-4-06; (Fotel & Thomsen, 2003)]

The *surveillance* collocates *airport* and *airports* represent sites of high mobility and can be understood as *filters* to this mobility as suggested in Example (20) from Issue 1(4). Surveillance in this context is particularly linked to the concept of security, reflected in the co-occurrence

pair (*security, airport*) occurring significantly more frequently in Volume 1 compared to the full corpus (see Figure 4-3).

- (20) This paper illustrates the surveillant sorting that is perhaps most illustrative of airport **surveillance**, where *airports* can be seen to act as filters (Lyon, 2003) to the mobilities that pass through them.

[2003-01-4-04; (Adey, 2003)]

In Mehrabov's (2015) content analysis the final theme "Work, Power and Resistance" is matched to Issue 1(2) of Volume 1, which is simply titled "Work". How surveillance scholars understand surveillance in the work context is exemplified in (21). This example offers a specific paraphrase of *workplace surveillance*, a common cluster in the concordance of the co-occurrence pair (*surveillance, workplace*) from Figure 4-2. This paraphrase mirrors Lyon's broad surveillance definitions, especially in its use of the phrase *for the purposes of*; cf. Example (6), Section 4.2.21.

- (21) When referring to *workplace* **surveillance** I mean the multiplicity of formal and informal practices of monitoring and recording aspects of an individual or groups' behaviour 'at work' for the purposes of judging these as appropriate or inappropriate; as productive or unproductive; as desirable or undesirable ; and so forth.

[2003-01-2-06; (Introna, 2003)]

Another article in the "Work" issue illustrates a specific theoretical approach called "social exchange theory" for examining such surveillance practices at the workplace and their privacy implications (2003-01-2-03; Stanton & Stam, 2003). The discussion of this theoretical framework is reflected in the co-occurrence pair (*social, exchange*) in Figure 4-3.

The two collocates of *surveillance* on the right side of Figure 4-2 lack effect size boxes, indicating that *surveillance* only co-occurs with *health* and *scholars* in the later volumes (and

does not co-occur at all with them in Volume 1). One of the reasons that *health* is listed as a significantly more frequent collocate in the full corpus is that it is a particularly salient collocate in Volume 6, which contains a special issue on surveillance in the area of health and medicine. I return to this topic in the next subsection.

4.4.2.2 Volume 6 (2009) – The relaunch

Volume 6 consists of four special issues, so all of them have been allocated to themes by Mehrabov (2015), covering all themes with the exception of “Mobility and Stasis”, which was very salient in the results of Volume 1. As with the previous section, I go through the themes and reflect on them in relation to the co-occurrence results for Volume 6 compared to the full corpus (Figures 4-4 and 4-5).

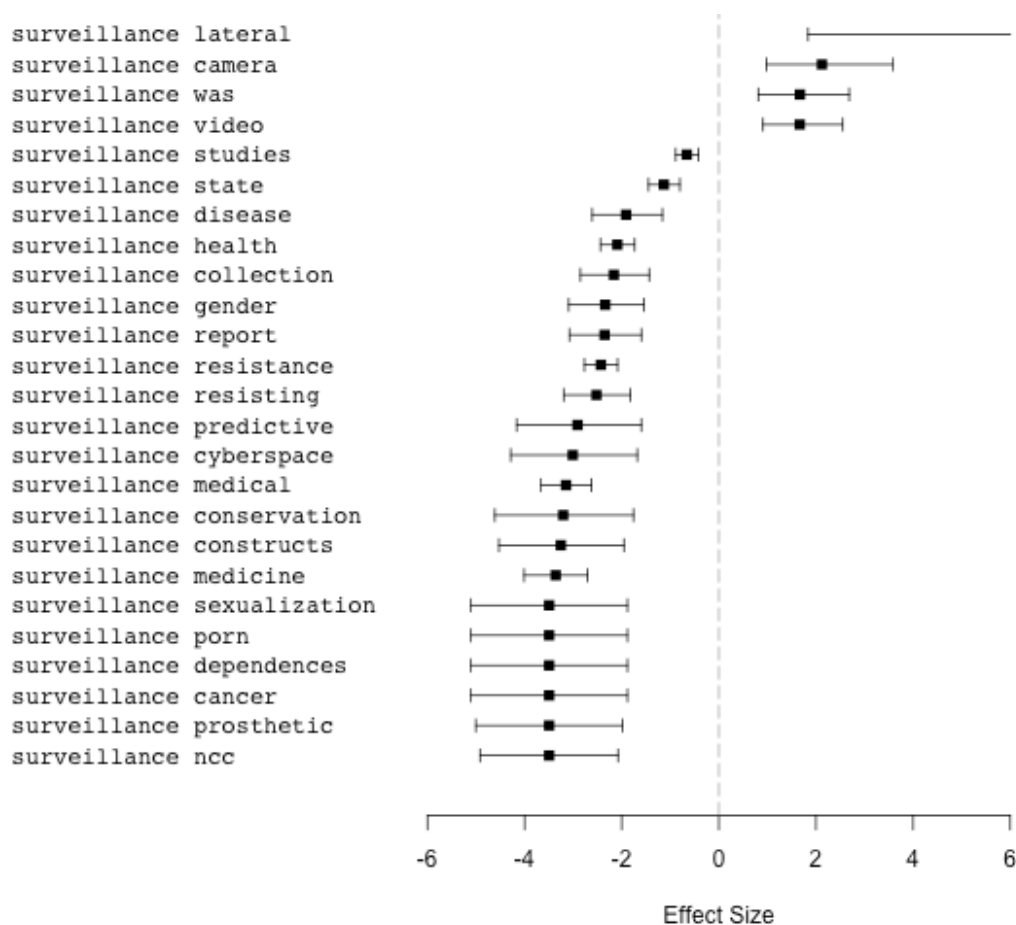


Figure 4-4: Co-occurrences of surveillance in Volume 6 vs. the whole S&S Corpus

Mehrabov (2015) allocates Issue 6(1) to the first theme, “Classic Surveillance”. This issue marks a milestone in the publication history of the S&S journal, as with this issue the journal was relaunched on a new website, hence the title “Relaunch Issue: Revisiting Video Surveillance”. The following quote shows the beginning of a statement that was posted with the issue on the website together with the relaunch issue:

This issue is the first on our new website, powered by the Public Knowledge Project’s Open Journal System. This deepens Surveillance & Society’s commitment to Open Source and Open Access.

This Relaunch Issue not only revisits one of the key contemporary technologies of surveillance, CCTV, by placing it in deeper historical context, but also reconsiders the past, present and future of Surveillance Studies.

(The Editors, 2009)

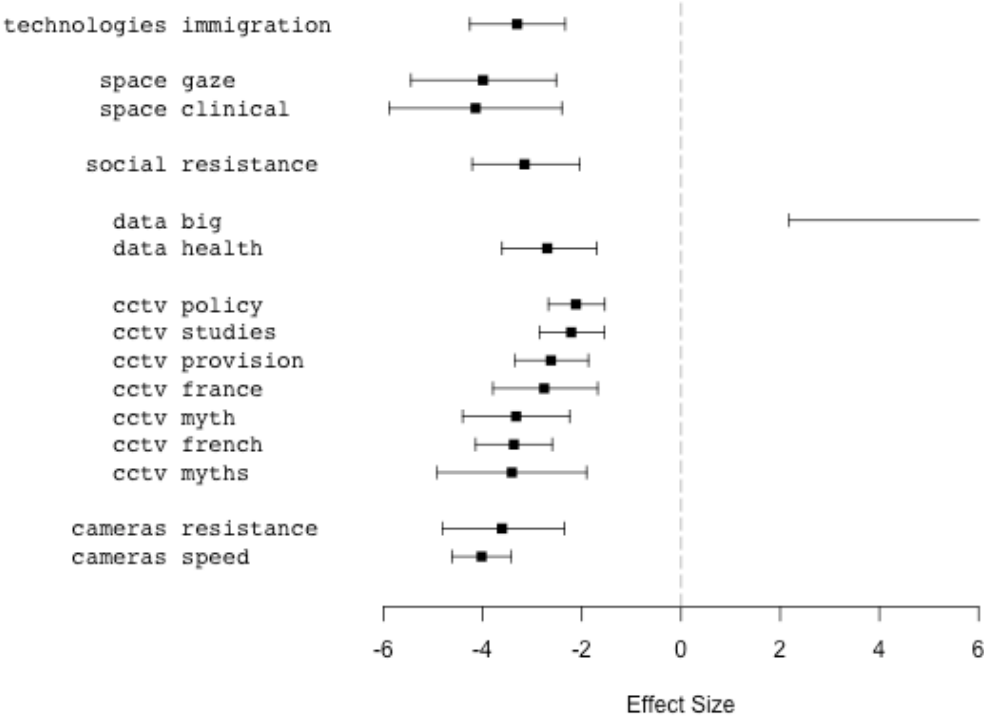


Figure 4-5: Co-occurrences of other KKWs in Volume 6 vs. the whole S&S Corpus

The editorial note refers to two objectives of the issue; “revisiting” CCTV and “reconsidering” the development of surveillance studies. The *re-* prefixes in both verbs indicate a reflective approach in Issue 6(1). This suggests a contribution to the development of the field, even though Mehrabov (2015) has categorised this issue among “classic” surveillance studies.

The editors describe CCTV as “one of the key contemporary technologies of surveillance”. Matching this emphasis on CCTV on the website and the associated mention of *video surveillance* in the title of the special issue, the co-occurrence comparison finds several collocates of the node *CCTV* to be significantly more frequent in Volume 6 compared to the full corpus (see Figure 4-5). The most distinct collocate of *CCTV* in Volume 6 is *myths*. Together with its singular form, it is found in an article that explores so-called “CCTV myths” for which little evidence exists: (i) “CCTV works”; (ii) “CCTV is everywhere”; (iii) “Citizens want CCTV”; (iv) “Citizens understand the technological capabilities of CCTV”; and, (v) “CCTV is there to protect us and reduce crime” (2009-06-1-03; W. Webster, 2009).

Another article from Issue 6(1) focuses on the tradition of “French CCTV studies” (see the *CCTV* collocates *French*, *studies* and *France* in Figure 4-5). This article exemplifies the observations from Section 4.2.1 on the interdisciplinarity and transnationalisation of surveillance studies. It highlights the potential challenges of this diversity, especially when research is disseminated in different languages, and therefore encourages more engagement with the “‘lost’ CCTV studies published in French academia” (2009-06-1-04; Klauser, 2009).

Beyond the focus on CCTV surveillance, Issue 6(1) contains more general reflections on the nature of surveillance studies. A combined review of a reader and an introductory textbook provides salient examples of this self-reflection. The publication of these discipline-forming publications is presented as a sign of the discipline “coming of age”. Nevertheless, the article expresses concerns that the field might become too “institutionalised”; see Example (22). It

strongly argues for surveillance studies as a concept to stay open to “be debated and discussed”, in a similar vein to Teubert’s (2010, p. 269) argument that “[a] healthy discourse is plurivocal”.

(22) I am wary of Surveillance Studies, as it becomes more recognized, becoming at the same time more institutionalised. However the scope and direction of the field needs to be debated and discussed.

[2009-06-1-08; (Murakami Wood, 2009b)]

To the second theme, “Identity-based Surveillance”, Mehrabov (2015) allocates two issues, 6(2) on surveillance in the context of health and 6(4) on surveillance and gender. Many of the distinct collocates of *surveillance* in Figure 4-4 relate to this identity theme. The identity-based topics of the special issues are interrelated. While some specific collocates only occur in individual articles, others are shared across the two special issues. The collocates with the highest effect sizes (at the bottom-left corner of Figure 4-4), from *sexualization* to *NCC*, are refined to individual articles and therefore rather specialised (see the wide confidence intervals).

Two more widely used collocates of *surveillance* are *medical* and *medicine*, most commonly forming the bigrams *medical surveillance* and *surveillance medicine*. There are some similarities in the way the terms are used (e.g. both occur in different *forms*; lines 16–17 in Concordance 4-12, lines 17–18 in Concordance 4-13). Overall, however, the co-occurrence patterns differ. *Medical surveillance* is preceded by *deployment* (Concordance 4-12, line 13), *limits* (lines 20–21), *manipulations* (line 22), *mechanisms* (line 23), *process* (line 27), *techniques* (lines 28–31) and *technologies* (lines 32–33). By contrast, *surveillance medicine* is described as *a concept* (Concordance 4-13, lines 12–13) that is *defined* and has been *called* (line 3) this *term* (line 26) based on its *specific characteristics* (line 11). A *defining attribute* can be listed, which in this case is *turning the gaze upon oneself* (extended context of line 9). Other issues can be found *at the centre of surveillance medicine* (line 10) and the concordance talks

about the *expansion* (line 15), *extension* (line 16) and *infiltration of medical surveillance into* (lines 19–20) new domains. In short, *medical surveillance* tends to refer to specific techniques (such as *cervical cancer screening*, line 10). It is not a defined concept, unlike *surveillance medicine*, which has been theoretically conceptualised.

1 er discussions about **medical surveillance** in the emerging conditions o 2009-06-2-07
2 edical discourse and **medical surveillance**. In this special edition of 2009-06-2-01
3 ered eating. Because **medical surveillance** - as an apparatus of power/k 2009-06-2-06
4 r, that contemporary **medical surveillance** also utilizes 'bottom-up' an 2009-06-2-01
5 ence of cyberspatial **medical surveillance** must be seen in the context 2009-06-2-07
6 nge and the diffused **medical surveillance** that informs the censure of 2009-06-2-06
7 ly 'constituted' for **medical surveillance**. What could possibly make th 2009-06-4-05
8 for considering how **medical surveillance** takes place online. In some 2009-06-2-07
9 k whether increasing **medical surveillance** does, indeed, constitute a d 2009-06-2-01
10 cancer screening is **medical surveillance**; but given the intended outc 2009-06-4-05
11 lessly as non-trans. **Medical surveillance** focuses first on individuals 2009-06-4-02
12 urce of criticism of **medical surveillance**, drawing attention to the di 2009-06-4-05
13 le the deployment of **medical surveillance** a new, ubiquitous feature of 2009-06-2-07
14 ious expectations of **medical surveillance**. This infiltration of health 2009-06-2-07
15 their experiences of **medical surveillance** and authority. Pro-anorexia 2009-06-2-06
16 distributed form of **medical surveillance** can be seen to represent bot 2009-06-2-01
17 led to a new form of **medical surveillance**, which we refer to as prosth 2009-06-2-07
18 res two instances of **medical surveillance** that illustrate post-panopti 2009-06-2-03
19 ence on the issue of **medical surveillance** or clinical monitoring is li 2009-06-2-03
20 ia and the limits of **medical surveillance** over the (female) body. Bell 2009-06-2-01
21 ia and the limits of **medical surveillance**. Since it emerged into virtu 2009-06-2-06
22 xic manipulations of **medical surveillance** thus locate pro-anas in a li 2009-06-2-06
23 of the mechanisms of **medical surveillance**. The impact of computing and 2009-06-2-03
24 plex negotiations of **medical surveillance** undertaken by participants i 2009-06-2-06
25 plex negotiations of **medical surveillance** undertaken by those who are 2009-06-2-01
26 the pervasiveness of **medical surveillance**, Rich and Miah recognize tha 2009-06-2-01
27 a broader process of **medical surveillance** within cyberspace. Cyberspac 2009-06-2-07
28 employ techniques of **medical surveillance** in order to actively make se 2009-06-2-03
29 gured. Techniques of **medical surveillance** and monitoring, the authors 2009-06-2-01
30 w that techniques of **medical surveillance** now include more bottom-up a 2009-06-2-03
31 se the techniques of **medical surveillance** to make sense of their own b 2009-06-2-01
32 the technologies of **medical surveillance**, but is expected to engage i 2009-06-2-01
33 the technologies of **medical surveillance** in the 21st Century. e 2009-06-2-01
34 r Oneself: themes of **medical surveillance** in society Susanne Bauer Med 2009-06-2-03
35 illance & Society on **medical surveillance**. While Wald deals less direc 2009-06-2-08
36 are enormous. Online **medical surveillance** influences not just how indi 2009-06-2-01
37 nd wellbeing. Online **medical surveillance**, Rich and Miah argue, serve 2009-06-2-01
38 erson should require **medical surveillance**. Some residents with certain 2009-06-2-04
39 rough which to study **medical surveillance** in cyberspace due to their m 2009-06-2-07
40 authors suggest that **medical surveillance** has become more complex in r 2009-06-2-01
41 ic discourse and the **medical surveillance** of particular 'anorexic bodi 2009-06-2-06
42 ndEn 2001: 229). The **medical surveillance** of prosthetic bodies reflect 2009-06-2-07
43 edition explore the **medical surveillance** of the human body. The artic 2009-06-2-01
44 cradle to grave, the **medical surveillance** of the human body has, for m 2009-06-2-01

Concordance 4-12: All instances of *medical surveillance* in Volume 6

1 of what has been defined as 'surveillance medicine': This new surveill 2009-06-2-07
2 s new mechanism of power as 'surveillance medicine'. In what others ha 2009-06-2-01
3 med by what has been called 'Surveillance medicine' by David Armstrong 2009-06-4-05
4 ic extension'. Consequently, 'surveillance medicine', now reaches into a 2009-06-2-07
5 ch to articulate and extend 'surveillance medicine' (Armstrong 1995) a 2009-06-2-07
6 zation of cyberspace expands 'surveillance medicine' into spaces that we 2009-06-2-07
7 act This paper examines how 'surveillance medicine' (Armstrong 1995) h 2009-06-2-07
8 eillance medicine': This new 'surveillance medicine' involves a fundamen 2009-06-2-07
9 been a defining attribute of 'surveillance medicine'; now such self exam 2009-06-2-01
10 on studies. At the centre of 'surveillance medicine', in the sense the t 2009-06-2-03
11 specific characteristics of 'surveillance medicine' and its capacity to 2009-06-2-07
12 mstrong's (1985) concept of 'surveillance medicine', Emma Rich and And 2009-06-2-01
13 395). Armstrong's concept of 'surveillance medicine' draws attention to 2009-06-2-07
14 the constitutive effects of 'surveillance medicine' expands in the sens 2009-06-2-07
15 on makers'. The expansion of 'surveillance medicine' is, as noted above, 2009-06-2-07
16 boundaries and extension of 'surveillance medicine' in cyberspace. Howe 2009-06-2-07
17 linary spaces, a new form of 'surveillance medicine' emerged in the twen 2009-06-2-11
18 c bodies. Thus, this form of 'surveillance medicine', not only regulates 2009-06-2-07
19 examine the infiltration of 'surveillance medicine' into those cyberspa 2009-06-2-07
20 anced by the infiltration of 'surveillance medicine' into non medical ar 2009-06-2-07
21 h which to study the kind of 'surveillance medicine' Armstrong (1995) ou 2009-06-2-07
22 sort of moral narratives of 'surveillance medicine', which have traditi 2009-06-2-07
23 al spaces for the purpose of 'surveillance medicine'. The Nintendo Wii i 2009-06-2-07
24 ce is focused on the sort of 'surveillance medicine' described above, wh 2009-06-2-07
25 s future capacity to promote 'surveillance medicine'. As Cummins et al. 2009-06-2-07
26 d Armstrong coined the term 'surveillance medicine' as 'a significant 2009-06-2-03
27 nt ethical issues given that 'surveillance medicine' now targets increas 2009-06-2-07
28 tern society. Reflecting the 'surveillance medicine' outlined above, per 2009-06-2-07
29 of knowing the body through 'surveillance medicine'. Recent reactions f 2009-06-2-07
30 cts a broader shift towards 'surveillance medicine' (Armstrong 1995). 2009-06-2-07
31 ealness, illustrates the way 'surveillance medicine' is operating within 2009-06-2-07
32 berspace. To conclude, while 'surveillance medicine' regulates physical 2009-06-2-07

Concordance 4-13: All instances of *surveillance medicine* in Volume 6

In addition, the collocates *health* and *disease* are used to discuss surveillance in medical contexts. Example (23) suggests that *disease surveillance* does not focus on *individuals* and is therefore less personal than other types. *Disease surveillance* can also refer to animals, as the analysis of surveillance discourse in *The Times* shows (see Section 6.3.1).

- (23) the idea that **surveillance** applies to *disease*, rather than to individuals, distinguishes **surveillance**, for public *health* purposes, from control activities.
[2009-06-2-02; (French, 2009)]

The overlap between the special issue themes of health in 6(2) and gender in 6(4) is visible from Concordances 4-12 and 4-13. Medical surveillance and surveillance medicine occur mostly in the health issue, but are also mentioned several times in the gender and sexuality issue. An example is the surveillance technique of cervical cancer screening, which is framed

as “a fruitful and appropriately complex instance of the ambiguous mutual dependences of gender constructs and surveillance practices” (2009-06-4-05; Corones & Hardy, 2009).

The space that the surveillant gaze focuses on in the medical context differs from space as conceptualised in relation to mobility, where it refers to spaces like “shopping malls” or “neighbourhoods” (cf. Section 4.4.2.1). Example (24) illustrates an occurrence of the pair (*space*, *gaze*) from Figure 4-5. It shows that in the medical context, the main sight of surveillance is traditionally the “body”, but due to modern medical surveillance techniques the “gaze” is increasingly more “distributed”. The reference to *post-panoptic* further indicates that this conceptualisation is moving beyond the theme of “Classic Surveillance”.

(24) Surveillance techniques directed at individual patients and at population health reconfigure the constellation of the body, **space** and the *gaze* into a post-panoptic distributed mode.

[2009-06-2-03; (Bauer & Olsén, 2009)]

Despite the different perspectives on space, mobility and medical surveillance are not mutually exclusive. As a case in point, the co-occurrence pair (*technologies*, *immigration*) occurs in a case study of immigrant medical examinations. Both *immigration* and the *border technologies* mentioned in Example (25) can be viewed as forms of mobility-based surveillance.

(25) By evaluating technologies of immigrant medical examinations and attempts to liberalize and humanitarianize *immigration* and border **technologies**, this paper presents [...]

[2009-06-2-04; (Wiebe, 2009)]

The fourth theme, “Work, Power and Resistance” is matched with Issue 6(3) by Mehrabov (2015) due to its special issue title, “Surveillance and Resistance”. Example (26), showing the co-occurrence pair (*surveillance*, *resistance*) from Figure 4-4, illustrates the importance that

some theorists attribute to this topic. Here, resistance is framed as a “co-development of surveillance” rather than “merely an epiphenomenon”.

- (26) Just as surveillance has become a normalised part of everyday life, *resistance* to **surveillance** is equally ‘normal’ (De Certeau 2002). Yet resistance is not merely an epiphenomenon of surveillance - it is a basic and necessary co-development of surveillance [...]

[2009-06-3-03; (Martin et al., 2009)]

Further links can be found between the themes of “Work, Power and Resistance” and “Identity-based Surveillance”. Example (27) provides corpus evidence for explicit links between resistance and “issues of identity”, illustrating a particular type of resistance: “resistance to speed cameras” (see the collocates *resistance* and *speed* of the KKW *cameras* in Figure 4-5).

- (27) Issues of identity are central to most social and resistance movements (Buechler, 2000) and are certainly so in *resistance* to *speed* **cameras**.

[2009-06-3-05; (Wells & Wills, 2009)]

Less obvious collocates for the theme of resistance are *sexualization* and *porn*. Unexpectedly, they do not belong to Issue 6(4) on gender and sexuality. Instead, an article in the resistance issue, 6(3), argues that the contexts related to these collocates can contribute to resistance.

Figure 4-4 shows that three collocates of *surveillance* are underrepresented in Volume 6 (*video*, *was*, *camera*) and that *lateral* is a unique collocate in the full corpus. Due to the focus of Issue 6(1) on CCTV surveillance, it is unexpected to find *video* and *camera* underrepresented in this volume. Yet, the corpus data suggests that *CCTV* is the preferred term in Volume 6, as *CCTV* is over 10 times as frequent as *video* (405 vs. 40 total occurrences, not only as a collocate of *surveillance*) and *video surveillance* only occurs 11 times.

In summary, this section has given an overview of the collocates contributing to the themes of “Classic Surveillance”, “Identity-based Surveillance”, as well as the fourth theme, “Work, Power and Resistance” from Mehrabov’s (2015) content analysis. The discussion has shown that these themes are interrelated and not necessarily restricted to any single special issue. In addition, mobility surveillance related to the theme of “Mobility and Stasis” has been identified in the co-occurrence patterns, although none of the issues in Volume 6 have been formally allocated to this theme by Mehrabov (2015). The reason for this discrepancy is that Mehrabov’s (2015) content analysis focuses on a different level of meaning-making of the journal than the corpus linguistic approach. Through his focus on special issue topics, Mehrabov’s (2015) analysis is concerned with the meaning-making level of “Issues” of Figure 3-1 in Section 3.2. My corpus linguistic analysis works with both the higher level of volumes and the lowest level of lexical patterns, which can then be linked back to particular articles. So, even though none of the issues in Volume 6 fully focus on mobility, that theme is part of the surveillance discourse – and the corpus approach is able to identify its prominence in particular articles of this Volume.

4.4.2.3 Volume 13 (2015): The final volume of the S&S Corpus

Volume 13 differs from the two issues discussed in the previous section, because it does not form part of Mehrabov’s (2015) content analysis and therefore has neither been allocated to any of his themes nor contributed to the formation of those themes. In fact, Mehrabov’s (2015) review article was published in issue 1 of this volume and is part of the corpus (file 2015-13-1-09). A second parameter by which Volume 13 differs from the previous two is that its representation in the corpus is incomplete: at the time of corpus compilation Issue 13(2) was the latest issue available – the double issue 13(3/4) was published after the data collection had

ended. Since the S&S journal is an ongoing publication, the corpus was always going to be incomplete and just presents a snapshot of this journal. Nevertheless, the corpus provides a wealth of data about the field of surveillance studies, spanning the first fourteen years of the journal's publication history.

The first issue of Volume 13 focuses on the scholarly field: with its title "Doing Surveillance Studies (2)" it is framed as a sequel to Issue 3(1/2), "Doing Surveillance Studies". Mehrabov (2015) did not allocate that first surveillance studies issue to any theme. The relaunch issue 6(1), which also has a self-reflective aim (see Section 4.4.2.2), was allocated to the "Classic Surveillance" theme (Mehrabov, 2015) owing to its focus on CCTV Surveillance. The results from Volume 13 (Figures 4-6 and 4-7) do not suggest any immediate link to "Classic Surveillance" for this last volume. On the contrary, the co-occurrence patterns point to a number of innovative trends.

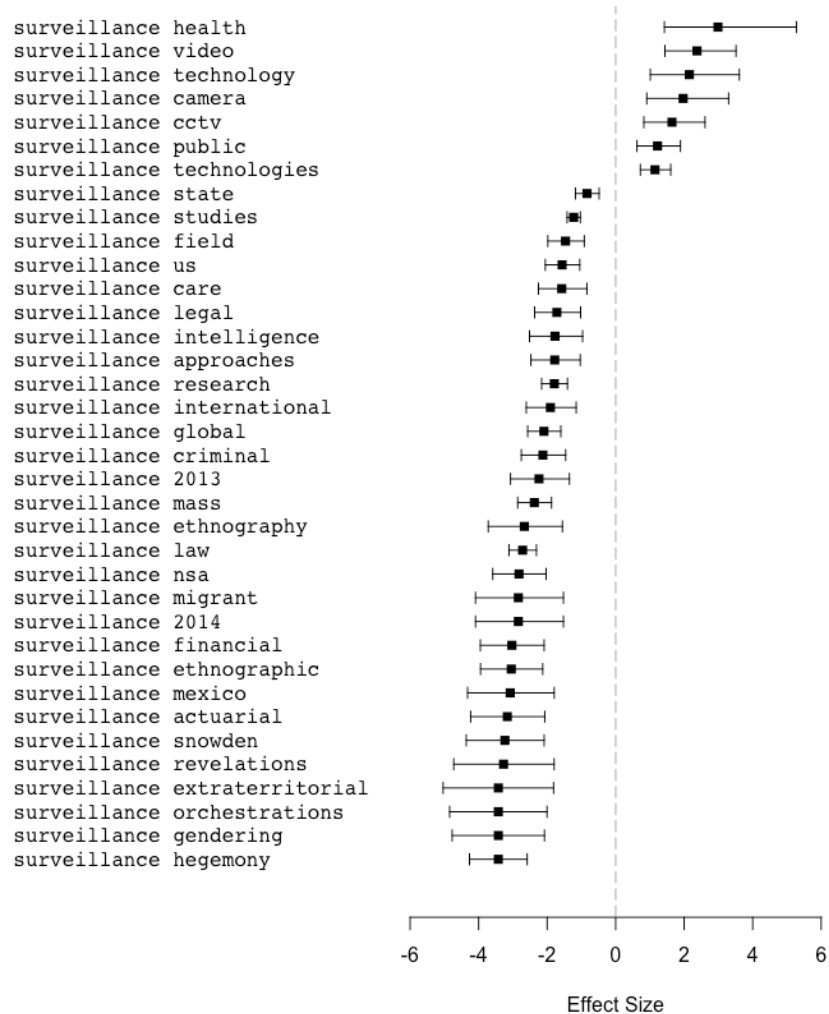


Figure 4-6: Co-occurrences of *surveillance* in Volume 13 vs. the whole S&S Corpus

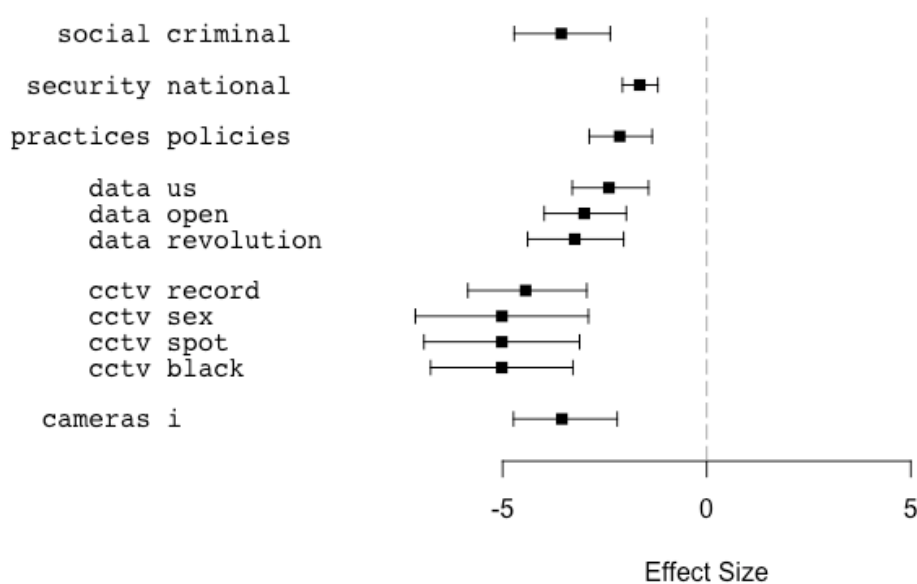


Figure 4-7: Co-occurrences of other KKWs in Volume 13 vs. the whole S&S Corpus

The issue's focus on self-reflexivity on surveillance studies can be seen in concordance lines for the *surveillance* collocates *ethnographic* and *care*. Example (28), from the article in which *ethnographic* occurs most frequently with *surveillance*, illustrates the scholars' proposals for applying *ethnographic practice* more transparently. Crucially, the article criticises the "unclear" use of *surveillance*, underlining the importance of surveillance definitions in this community (see Section 4.2.2).

(28) Despite the widespread acceptance and practice of ethnographic strategies in Surveillance Studies however, some things remain very unclear in empirical ethnographic accounts, not least amongst them the deployment of the term 'surveillance' itself.

[2015-13-1-03; (Green & Zurawski, 2015)]

Care is particularly frequent in another article from Issue 13(1), which differentiates the concepts of "Surveillance as Care vs. Surveillance as Control". Example (29) illustrates the article's argument about the theme of care acting as a balance for control, strikingly echoing Example (19) (cf. Section 4.4.2.1), published 12 years earlier: "surveillance has two faces and is a matter of both care and control" (2003-01-4-06; Fotel & Thomsen, 2003). This observation suggests that change can be slow and not all contributions to the discourse are equally influential (see Teubert, 2010).

(29) Placed in relation to **Surveillance** Studies, an emphasis on "*care*" may also temper the dominance of themes of "control" that characterize the field, [...]

[2015-13-1-04; (Abu-Laban, 2015)]

The second issue of the volume, 13(2) is titled "Surveillance and Security Intelligence After Snowden (Part 1)" and relates to the series of leaks of documents from the US National Security Agency (NSA) by Edward Snowden, in 2013. The collocates related to this issue exhibit a

semantic field of a wide scope that was not present in the results for other volumes. In a sense these new meanings of surveillance build onto the earlier observations on mobility, such as Example (17), Section 4.4.2.1, which suggests that an individual's activities at different physical places are connected via a "flow of surveillance". The argument that surveillance is widespread has been made in previous volumes and is also reflected in the wide range of surveillance case studies that have been discussed in this chapter. However, Volume 13 provides the first set of co-occurrence results, compared to the full corpus, that indicate the spread and scope of surveillance as a salient pattern (see Figure 4-6). These collocates emphasise both the geographic reach (*extraterritorial, global, international*) and the large scale of surveillance measures (*mass*; also see Concordance 4-14, line 4) disclosed by the released documents.

The *surveillance* collocate with the largest difference is *hegemony*. It only co-occurs with *surveillance* in one article of the corpus and is used in the political context of the NSA documents. The article focuses on political deals that the US have with other states in order to facilitate surveillance, distinguishing those more *public* measures from the *secretive programs* mentioned in (30).

- (30) The National Security Agency activity disclosed by Edward Snowden plugs into a larger information ecology made possible by US **surveillance** *hegemony*. While the revelations of the NSA's international spying ambitions have astonished, there is more to US surveillance than secretive programs carried out by its intelligence community.
[2015-13-2-04; (Keiber, 2015)]

1 d Lyon, author of a new book **surveillance** after the Snowden revelation 2015-13-2-01
2 tions have revealed that the **surveillance** agencies concerned are, inde 2015-13-2-11
3 the Snowden revelations for **surveillance** and security intelligence an 2015-13-2-01
4 s the revelations about mass **surveillance** continue. As Ron Deibert ind 2015-13-2-02
5 s? The revelations about NSA **surveillance** have obviously shaken up the 2015-13-2-05
6 ecology made possible by US **surveillance** hegemony. While the revelati 2015-13-2-04
7 rgued here that the kinds of **surveillance** highlighted by the Snowden r 2015-13-2-02
8 velations have evidenced NSA **surveillance** of corporations such as Petr 2015-13-2-03
9 evelations indicate that NSA **surveillance** of the entire mobile telepho 2015-13-2-03
10 is increasingly subjected to **surveillance**, particularly salient follow 2014-12-4-03

Concordance 4-14: All instances of (*surveillance*, *revelations*) in the S&S Corpus

At the same time, the example concedes that the disclosures “have astonished”. Here and elsewhere, Snowden’s actions are described as *revelations*. Concordance 4-14 lists all ten instances. Like *hegemony*, *revelations* only co-occurs with *surveillance* in Issue 13(2), though, as the concordance shows, the pair co-occur in several articles. *Revelations* was used commonly in the newspaper coverage of the Snowden leaks. In their keyword analysis of different newspapers covering the events, Branum and Charteris-Black (2015, p. 215) find the word *revelations* to be key in *The Guardian* compared to a reference corpus of various newspapers. They argue that *revelations* has “religious connotations” and a “positive semantic prosody” (Branum & Charteris-Black, 2015, p. 215). By contrast, *leaks* is the keyword used by the tabloid *The Sun* (Branum & Charteris-Black, 2015, p. 215). Since *The Guardian* published the initial leaks, it may not be surprising that it presented the information as *revelations*, suggesting novelty and magnitude. However, the same term was used, for example, in the popular books *No Place to Hide* (Greenwald, 2014/2015) and *The Snowden Files* (Harding, 2014).

The use of this word by surveillance studies scholars appears to lend more authority to the claims about the importance of the implications in the disclosures for public knowledge about surveillance. The corpus even contains a suggestion that the *revelations* have “shaken up” the field (2015-13-2-05; van der Velden, 2015). The same article explains that, as a result of the disclosures, “intelligence practices” have become “highly visible” to surveillance studies scholars beyond specialists in intelligence studies. The question following that statement,

shown in (31), illustrates the progress of meaning negotiation in the discourse of surveillance studies. The article goes on to propose a new framework for the surveillant devices and tools highlighted by the leaks and encourages more research on resistance developments in reaction to the “new reality of surveillance”.

(31) The question becomes: in what way are notions of rhizomatic, and less hierarchal, networks (Haggerty and Ericson 2000) and of self-exposure and ‘participative surveillance’ (Albrechtslund 2008) sufficiently able to make sense of this pile of knowledge about technologies of intelligence?

[2015-13-2-05; (van der Velden, 2015)]

The editorial is careful not to overstate the magnitude of the incident, describing the work in this S&S special issue as “impromptu historical periodizations”; see Example (32). The editors suggest that ongoing “revelations and discussions” will further clarify in how far the disclosed information is part of previous developments (2015-13-2-01; Murakami Wood & Wright, 2015).

(32) And, as with all such impromptu historical periodizations, there is always also a case to be made to say that Snowden’s revelations didn’t change as much as we thought it might, or at least, rested on a legacy of former events or long-standing processes.

[2015-13-2-01; (Murakami Wood & Wright, 2015)]

The NSA examples demonstrate how current affairs have an impact on surveillance discourses, not only in the popular media but also in the journal. Both (31) and (32) show that the process of researching the implications of the documents on theoretical frameworks of surveillance is ongoing at this stage of the publication history of the S&S journal. The Snowden revelations therefore presents a particularly good example of an evolving discourse. In this sense, patterns in corpora reflect social developments (also see Wiegand & Mahlberg, 2019).

4.5 Co-occurrence patterns across the corpus

In this section I show how the localised, volume-specific results identified in the previous section fit into the wider discursive patterns across the corpus. In the first subsection (4.5.1), I look for similarities across the volume-based results and present a larger-scale co-occurrence comparison of the journal volumes before and after its relaunch in 2009. Section 4.5.2 considers these patterns from a different perspective by proposing that they form a network of the surveillance discourses. The main focus is on the node *surveillance* and its salient collocates across the volumes, some of which are also KKWs.

4.5.1 Co-occurrence links and shifts across the volumes

This section argues that any collocate that appears in multiple sets of results for the volume-based co-occurrence comparisons has an important role in the discourse. Since the volume-based results are by definition salient in a given volume compared to the full corpus, overlap between these results must be limited. As seen in Section 4.4.2, the results for Volumes 1, 6 and 13 appear rather distinct. Going back to the co-occurrence comparison plots for *surveillance* in these Volumes (Figures 4-2, 4-4 and 4-6), reveals that only two collocates appear in the results for more than one volume – *state* and *studies*, both in Volumes 6 and 13.

Given that the plots are so specific, I argue that those collocates that are shared between these sets of results can point to special meaning components of *surveillance*. Table 4-8 provides a summary of these “shared, salient collocates”, i.e. collocates that appear as volume-specific collocates on more than one of the volume plots. (Henceforth, I use the shorthand “shared collocates”.) These collocates relate to the idea of consistent collocates (see Section 2.4.4; Gabrielatos & Baker, 2008), which are, however, originally based on collocation association measures. In Section 4.5.2 I discuss further how the relationships between these

shared collocates relate to collocation networks and connected concepts. As indicated by the double asterisks in Table 4-8, six of the 19 shared collocates (so almost a third) are KKWs: *cameras*, *CCTV*, *data*, *security*, *social* and *technologies*. This finding shows that these words do not achieve that “key” status across all volumes of the corpus because of their frequency independently of the surveillance discourse, but actively contribute to it. They are important collocates of *surveillance* and therefore shape its meaning in particular contexts that are reflected on the individual meaning-making levels below the full corpus – i.e. from the volumes, issues, individual articles down to the actual lexical patterns (see Figure 3-1, Section 3.2). The relative frequencies provided in the cells (per 100,000 words, in the corpus as a whole rather than the co-occurrence frequency with *surveillance*) show that even most of the words in Table 4-8 that are not KKWs are relatively frequent throughout the corpus. The only collocates that do not occur at all in one or more volumes of the corpus are *participatory* and *lateral*. These collocates point to more specific meaning constituents of *surveillance*.

Table 4-8: Shared collocates across the volume comparisons*

Collocate	Relative frequency (per 100,000 words) by volume*												
	1	2	3	4	5	6	7	8	9	10	11	12	13
children's	45.15	0.81	3.82	0.76	0.00	0.48	103.88	1.79	11.61	10.99	1.54	25.59	3.91
mobility	100.80	2.03	6.69	0.76	109.32	12.01	3.56	9.67	35.27	7.15	4.63	11.94	12.39
system	112.88	193.45	116.64	226.59	116.05	68.23	108.86	68.38	166.95	110.49	116.44	108.51	65.20
CCTV**	135.98	548.32	72.66	556.19	221.63	194.59	66.88	81.98	37.50	105.54	151.14	19.11	86.71
cameras**	91.35	186.56	35.37	147.25	124.29	59.58	69.73	73.39	25.00	223.18	50.90	25.93	36.51
video	37.28	126.13	64.05	37.38	17.97	19.22	45.54	78.05	8.04	76.96	40.10	43.68	13.69
technologies**	102.90	58.40	128.11	137.33	209.65	135.97	130.92	130.31	139.72	172.06	117.21	130.01	93.23
systems	142.80	211.30	67.88	263.22	134.77	98.50	57.63	79.12	134.81	87.95	101.40	96.57	85.41
border	5.25	5.27	4.78	9.92	307.73	26.91	9.25	87.00	36.16	29.68	15.81	7.85	8.48
security**	94.50	175.20	111.86	135.80	482.94	112.43	95.34	63.01	209.36	256.16	228.64	170.96	162.99
health	28.35	11.76	11.47	22.89	38.93	241.67	106.02	52.27	21.87	45.08	127.24	77.46	12.39
disease	15.23	1.62	0.96	1.53	11.98	57.18	9.25	3.22	2.23	1.65	14.27	34.12	0.65
state	60.90	96.52	145.32	64.85	186.44	182.10	119.54	107.76	158.47	122.58	101.79	113.63	177.33
studies	54.60	36.91	154.88	44.25	65.14	139.33	62.61	66.59	63.84	108.84	84.44	45.39	199.50
care	23.63	26.36	3.82	14.50	41.93	15.37	128.07	21.12	14.28	37.38	26.60	39.93	90.62
data**	276.16	143.57	266.74	144.20	204.41	162.88	108.86	332.23	111.15	217.13	158.08	528.92	299.25
participatory	1.05	1.62	1.91	0.00	0.75	1.92	1.42	53.70	4.91	11.54	12.72	18.77	4.56
social**	254.11	197.51	227.54	183.87	298.75	201.31	227.69	279.24	305.34	208.89	313.85	230.00	245.79
lateral	0.00	5.27	0.00	0.00	0.00	0.00	1.42	1.79	1.79	25.29	22.75	4.44	8.48
surveillance**	730.29	566.97	1,051.65	673.68	491.17	864.36	934.23	828.07	718.71	943.85	844.40	684.87	1,230.89

* Relative frequency in the corpus as a whole, not just the co-occurrence frequency with *surveillance*.

** The collocates marked with double asterisks (and *surveillance* itself) are also KKWs in the S&S Corpus (see Section 4.3).

Shaded cells in Table 4-8 specify the volumes in which a given collocate is significantly more frequent than in the full corpus. Among all the significant collocates that are shown on the *surveillance* plot (Figure 4-2, Section 4.4.2.1), only *children's* and *mobility* are highlighted in the column for Volume 1, because they are the only ones that also feature as significantly more frequent collocates in other volumes. The shaded cells show that, additionally, *children's* is a salient collocate of *surveillance* in Volume 7 and *mobility* in Volume 9. Hence, the shaded cells suggest similarities in co-occurrence between particular volumes of the S&S Corpus that can be considered as semantic links. Most collocates in Table 4-8 appear in the results for two volumes. The exception to this trend is the KKW *technologies*, which is featured in the results for three Volumes: 3, 5 and 10. This collocate appears to play a particularly important role in the meaning-making patterns of surveillance discourse.

The arrangement of the collocates visualises the temporal sequence in which the shared collocates co-occur saliently with *surveillance* in the subcorpora. So, the collocates have been ordered by the first volume in which they feature in the co-occurrence results. With their first appearance in the results of Volume 1, *children's* and *mobility* are shown at the top of the table. By contrast, *lateral* is displayed in the final row, because it first occurs in the significant results for Volume 10. Strikingly, *lateral*, like *participatory*, is (i) first found to be a salient collocate of *surveillance* towards the final volumes of the corpus and (ii) hardly occurs at all in the corpus prior to that, even outside the co-occurrence span of *surveillance*. Both collocates can be traced to conceptualisations of surveillance. Example (33) paraphrases *participatory surveillance* in terms of “an individual” and “known others” watching each other. Importantly, the “feeling of increased agency” that the “surveilled subject” is associated with suggests that the concept can be linked to the concept of “Empowerment” (see Mehrabov’s, 2015 third theme). The example explicitly associates *participatory* with *lateral surveillance*. The semantic link between the

collocates apparent in the example relates to the discussion in Section 4.2.2 on definitions. That is, the connection between the two concepts (in addition to a third one, *reciprocal surveillance*) in the example illustrates a meaning constituent of *surveillance* that is particularly pertinent to its use in the S&S Journal: the diverse but interrelated theoretical perspectives on surveillance. The discussion of these concepts and how they relate to one another is at the core of the meaning negotiation in the discourse of the S&S Journal.

- (33) Participatory surveillance is related to reciprocal surveillance (Regan and Steeves 2010) and lateral surveillance (Andrejevic 2005) and is distinct from the one-way top-down model of surveillance. In participatory surveillance, an individual watches known others and, in turn, those known others are watching the individual; being part of a participatory surveillance situation can give a feeling of increased agency for the surveilled subjects.

[2014-12-3-03; (Collister, 2014)]

The positions of *participatory* and *lateral* in Table 4-8 indicate that they represent relatively recent members of the shared *surveillance* collocates in this dataset. As the relaunch of the journal in 2009 is a milestone in its publication history (see Section 4.4.2.2), it is a useful point of reference for co-occurrence shifts over time. The double line between Volumes 5 and 6 represents this boundary between what I call “Part I” (Volumes 1–5) and “Part II” (Volumes 6–13) of the S&S Corpus. Based on this boundary, three groups can be distinguished. Group (i) contains collocates that make links only within Part I (i.e. they are salient across the early volumes). These collocates are listed in the upper half of Table 4-8: *system*, *CCTV*, *cameras*, *systems*. Group (ii) covers collocates that are more salient across volumes in Part II, shown in the lower half (*health*, *disease*, *state*, *studies*, *care*, *data*, *participatory*, *social*, *lateral*). Only six of the 19 shared collocates make links across the two parts: *children’s*, *technologies*, *border*, *security*, *mobility* and *video*. These collocates are allocated to group (iii). Because of their

saliency as collocates of *surveillance* across early and late volumes, they are argued to have a central role in the surveillance discourse.

More systematically, we can examine changes in the co-occurrence profile of *surveillance* in the earlier volumes (Part I) with that in the later volumes (Part II) by directly comparing the co-occurrence frequencies in these two large subcorpora. The full set of results for this comparison is shown in Figures 4-8 and 4-9. These figures show two halves of the same comparison, but for legibility are printed separately. Crucially, this wider comparison reveals that most of the collocates in the upper half of the table co-occur significantly more frequently with *surveillance* in Volumes 1–5 than Volumes 6–13: *mobility*, *system*, *CCTV*, *cameras*, *video* and *systems*. They are also underlined in Table 4-8, showing that most of the collocates that first occur as a salient collocate of surveillance in Part I are also overall more significantly frequent in this part. By contrast, italics in Table 4-8 mark collocates that are significantly more frequent in Part II in this direct comparison. As apparent from Table 4-8, that means that most of the collocates in the lower half of the table are significantly more frequent in Part II.

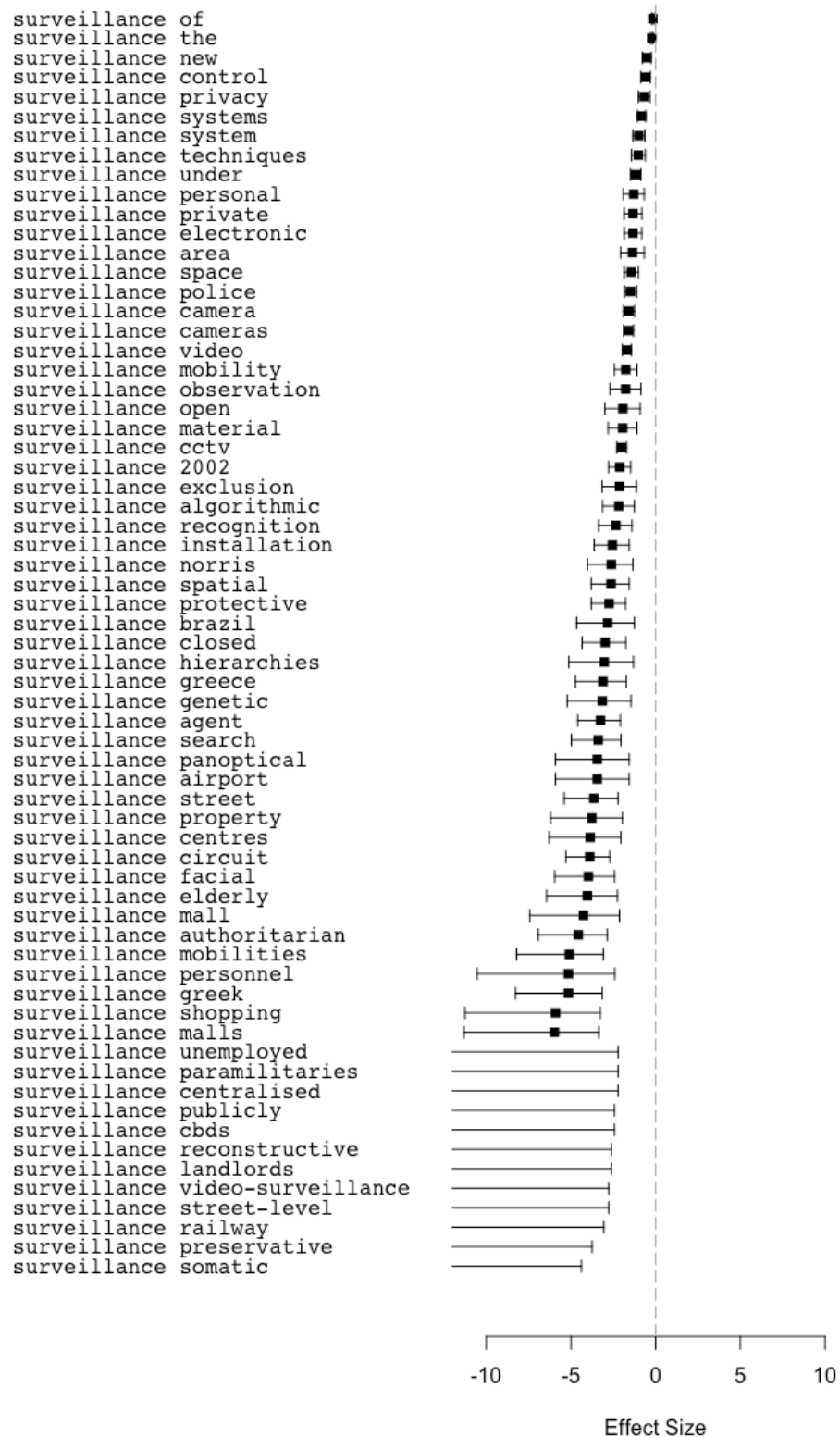


Figure 4-8: Significantly different co-occurrences of *surveillance* in Part I (Volumes 1–5) in comparison to Volumes 6–13

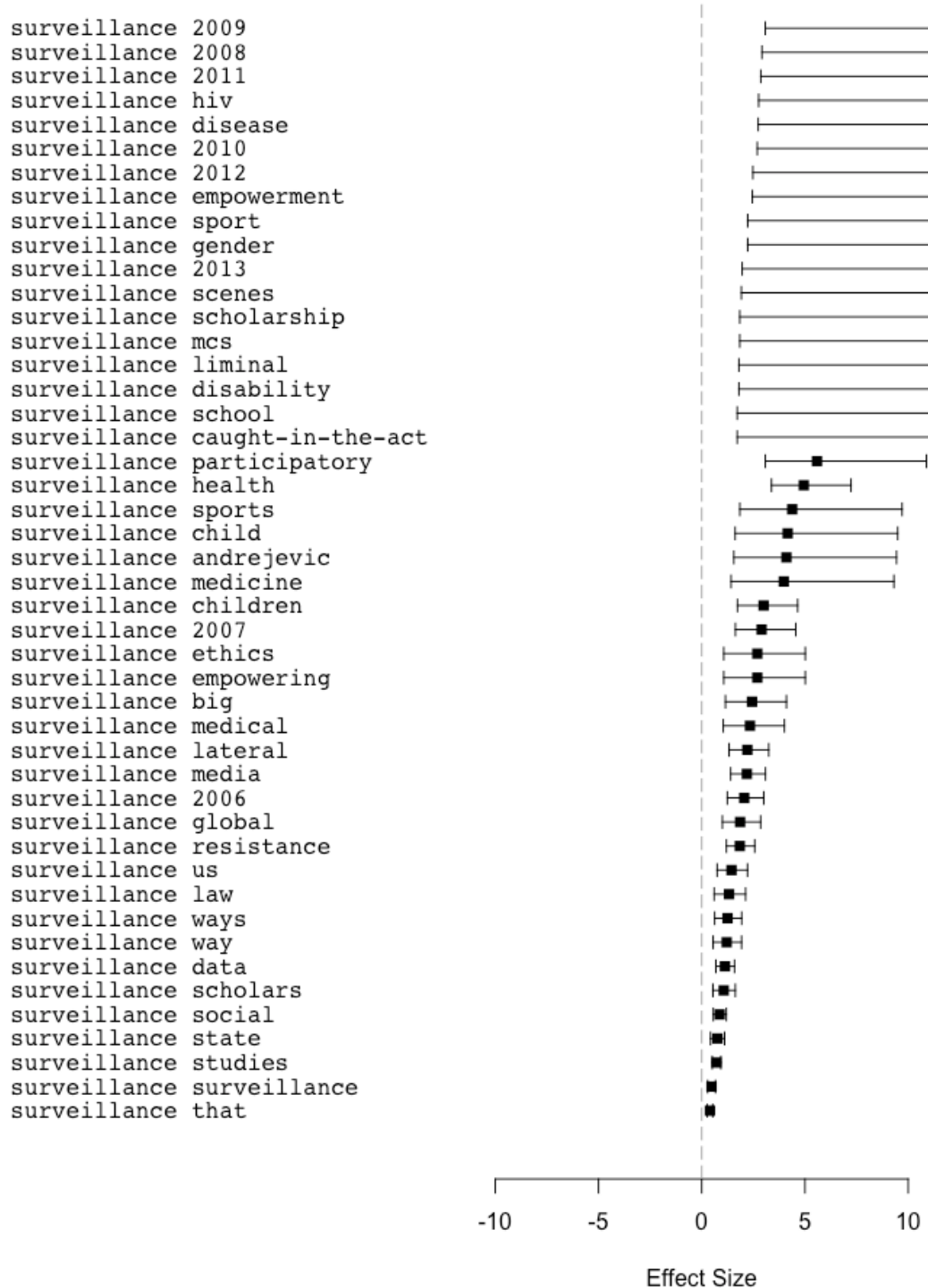


Figure 4-9: Significantly different co-occurrences of *surveillance* in Part II (Volumes 6–13) in comparison to Part I (Volumes 1–5)

The finding that the comparison plots match the impressions from the shared collocate plot supports the notion that there is a subtle change in the core discourses of surveillance across the

corpus. Yet this change is not extreme: as would be expected, the differences between Part I and II for the “shared” collocates from Table 4-8 are, overall, not as striking as for some really localised collocates that are not shared. That is to say, most of the collocates from Table 4-8 are located towards the centre line of the two plots, while the collocates with the highest differences are more distinct.

The unique collocates provide particularly striking examples of this distinctness, as some point to very specific types of change. Examples of unique surveillance collocates for Part I point to the concepts of *preservative* and *somatic surveillance*. The concept of *preservative surveillance*, which “aims to preserve public order and to prevent ‘anti-social’ behaviour” (2004-02-3_3-03; Klauser, 2004), occurs only in one article (in Volume 2) and is therefore not referred to intertextually in the S&S Corpus. *Somatic surveillance* relates to “the increasingly invasive technological monitoring of and intervention into body functions” (2007-04-3-01; Monahan & Wall, 2007). This concept is coined and discussed at length in one article of Volume 4. The term is used in a second article (in Volume 5) that refers back to the original article, so there is one intertextual reference. The overall lack of intertextual references to interpretations of *surveillance* as *preservative* or *somatic* therefore indicates a limited uptake of these concepts. Generally, unique collocates in Part I represent either a decrease in the interest in a particular theme or concept, or simply isolated contributions to the discourse. Only traces left in the discourse can indicate acceptance by the discourse community (see Section 4.2.2; Teubert, 2010, pp. 3, 115).

A different type of change is suggested by collocates that emerge in Part II. In its simplest, this change simply reflects the course of time: the individual years of the timespan 2008–2013 (2008, etc.) are all revealed as unique collocates of *surveillance* in Part II (see Figure 4-9). These years are mostly used for in-text citations published after the texts in Part I, reflecting

the temporal gap. They also show that field is evolving and the body of research being cited expands over time. Among the content words are, in a similar manner to Part I, collocates that refer to distinct concepts or themes. An example is *liminal*, which refers to a specific form of surveillance that is restricted to a “particular social event”, as illustrated in (34). The term is introduced in one article that proposes this term. Unlike with the unique collocates of Part I, the corpus evidence is not sufficient to conclude that this term has not been successfully taken up in the short period from the article’s publication in 2013 to the final volume in early 2015.

- (34) Liminal surveillance is the targeted, temporarily intensified use of a surveillance system for the safety and security management of a particular social event.
[2013-11-1_2-08; (Boersma, 2013)]

The unique collocates *empowerment* and *disease* differ from *liminal surveillance* in that they link to the wider themes of surveillance in the medical context or in relation to agency, power and resistance, also captured by Mehrabov’s (2015) themes. The related collocates *resistance* and *health* are both significantly more frequent in Part II (but not unique). Accordingly, the comparison of surveillance co-occurrence frequencies in Part I with Part II supports the findings about the shared collocates from Table 4-8 with additional evidence. In the following section I take a closer look at the links between the collocates in the table and their role in the surveillance discourse of the corpus as a whole.

4.5.2 A network of surveillance collocates in the S&S Corpus

This section is concerned with the links between the salient collocates of *surveillance*. It proposes that a network plot of these salient surveillance collocates can map out the main meaning constituents of *surveillance* in the S&S Corpus. I first explain the network plot of the

shared collocates in relation to the additional insights that can be gained from this way of displaying the data. Then I link the network of the surveillance collocates to the wider body of research on collocation and keyword networks and the concept of aboutness.

Figure 4-10 presents the network view of the data from Table 4-8. Each of the shared collocates forms a ‘node’ in the network.¹¹ The lines (also called ‘edges’) between the nodes represent the subcorpora (i.e. the volumes) in which two connected collocates both saliently co-occur with *surveillance*. The number of links from a given node is not only visible from the edges between the volumes, but is also reflected in the size of the node circles – the larger the circle, the more links it has.

Figure 4-10 emphasises relationships between the collocates that are not immediately obvious from the tabular perspective. The three thicker lines indicate that the connected words appear as these shared collocates in various volumes; each of the three pairs with thicker are listed together in two volumes:

- (*system, CCTV*) in Volumes 2 and 4;
- (*studies, state*) in Volumes 6 and 13; and
- (*data, participatory*) in Volumes 8 and 12.

All other links just indicate one shared volume (where the collocates are both significantly more frequent than in the full corpus). The nature of the co-occurrence comparison implies an emphasis of difference (in a similar way to the “difference bias” of regular key words mentioned in the literature review of keyness in Section 2.4.5). The fact that similarities are found across some volumes suggests that these collocates and their patterns contribute to the meaning-making constituents of *surveillance*.

¹¹ ‘Node’ is used in this section in the sense of “central points” in a network graph rather than node words in a concordance.

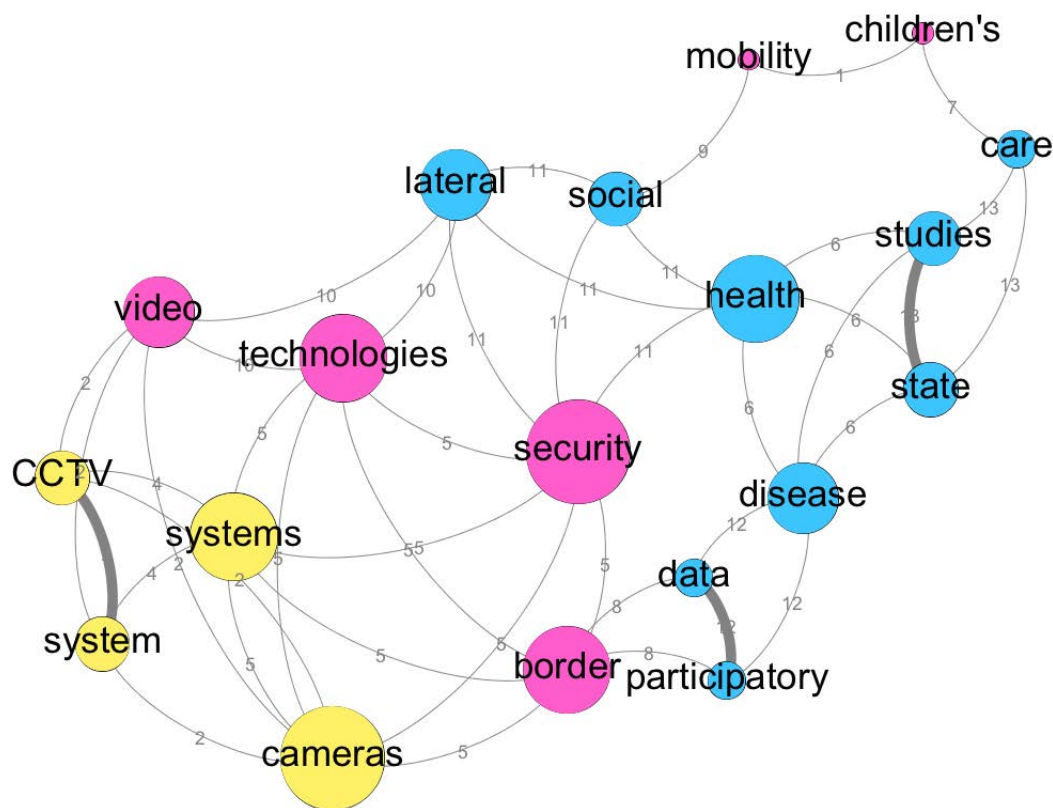


Figure 4-10: Volume links between the “shared collocates” created with *Gephi* version 0.9.2¹²

Group (i) of the shared collocates only form links within Part I of the journal (see Section 4.5.1). They are shown in yellow. The size of the nodes *cameras* and *systems* show that these collocates are particularly well connected among the early volumes. Given the apparent relationship to cameras and CCTV surveillance, this group appears to most closely relate to Mehrabov’s (2015) theme of “Classic Surveillance”. The blue nodes represent group (ii), i.e. those collocates that co-occur with *surveillance* at an outstanding frequency in later volumes. Within this group, *health* has the most links. Like *health*, several additional collocates of this group appear to relate to the theme of “Identity-based Surveillance” (Mehrabov, 2015): *care*, *disease*

¹² Layout created with the *Gephi* (Bastian, Heymann, & Jacomy, 2009) “Yifan Hu, “Noverlap” and “Expansion” algorithms.

and *social*. By contrast, the concepts of *lateral* and *participatory surveillance*, relate to ideas of surveillance (or resistance to it) as a means of empowerment (see Section 4.5.1).

Finally, group (iii), the collocates that link Parts I and II, are displayed in pink. *Security*, *border* and *security* from these group are located at a particularly central position, because they are among the nodes with the most links between volumes. Their high connectivity suggests that these nodes co-occur with *surveillance* frequently and in multiple contexts across the corpus. Two of these six collocates linking Parts I and II explicitly relate to Mehrabov's (2015) theme of "Mobility and Stasis" (*border*, *mobility*), suggesting that this theme is comparatively "timeless" in the surveillance discourse of the S&S Corpus. In this respect the mobility theme differs from the other three, as the shared collocates indicate that "Classic Surveillance" dominates in the early volumes, while the themes of "Identity-based Surveillance" and "Work, Power and Resistance" are salient in Part II.

These observations for the 13 volumes therefore resonate with the trends indicated in Mehrabov's (2015) allocation for special issues in Volumes 1–9 (see Section 4.4.1, Table 4-6). One development that cannot be clearly allocated to Mehrabov's (2015) theme is the shared collocate *state*, which shares links within Part II. This collocate seems to relate to the *mass surveillance* (mainly) by state-level organizations, as discussed in relation to the Snowden "revelations" in Section 4.4.2.3. This apparently rising attention on the state may indicate a new phase of "Classic Surveillance". As the methods seem to be different from the original "discipline and control" techniques and operate at a larger scale, an additional theme of "Mass Surveillance" might be warranted. This theme could also account for technological developments across the publication period that are apparent from both the KKW group of technologies (see Table 4-5, Section 4.3.2) and the collocates of KKW's identified in the comparisons. For instance, the concept of *big data* is only mentioned from Volume 10 onwards

and *big* is a significantly more frequent collocate of *data* in Volume 12 (see Appendix E). A method of testing the proposed theme of “Mass Surveillance” is to extend the S&S Corpus in order to trace the development of the co-occurrence network across recent volumes.

In a similar way to the KKW meaning groups, the network of the shared *surveillance* collocates sketches the aboutness of the corpus and by extension the main topics of relevance in the surveillance discourse. This network presentation relates to several previously proposed approaches to capturing meaning relationships and aboutness in corpora. Of the most obvious relevance is the recent work on collocation networks (Brezina et al., 2015; P. Baker, 2016). Brezina et al. (2015, p. 165) put forward the concept of ‘connectivity’ as an important component of collocation in addition to other characteristics previously identified in the literature (e.g. by Gries, 2013; see Section 2.4.4). They argue that collocation networks make it possible to discern “meaningful patterns” not only in the “narrow scope” of the span for surface co-occurrence, but “at the level of the text or discourse” (Brezina et al., 2015, p. 165). P. Baker’s (2016) study demonstrates how graph theory can be applied to collocation networks. He argues that different geometrical “shapes of collocation” can shed light on the relationships between words and, therefore, ultimately also on meaning (P. Baker, 2016, p. 139).

As suggested by Murakami et al. (2017; see Section 2.4.5), similarities exist between collocation networks and the machine learning technique of topic modelling, as both methods rely on co-occurrence counts. Given that topic modelling typically identifies the co-occurrences in larger stretches of text than the restricted span of usual collocation methods,¹³ Murakami et al. (2017) argue that the results of topic modelling point to more “thematic meanings” than collocation. Collocational meaning, they suggest, is more tied to phraseology.

¹³ This argument appears to focus on surface co-occurrence. Textual co-occurrence, as described by Evert (2008, p. 1222; see Section 2.4.4) would be able to consider co-occurrence relationships in paragraphs and entire texts.

The knowledge we can gain about meaning patterns across a corpus from association-based collocation networks also bears some resemblance to Scott's (1997; also see Scott & Tribble, 2006) networks of keyword associates. Associates of a keyword are other words of outstanding keyness in the same texts, which are therefore "co-key" with the keywords (Scott & Tribble, 2006, p. 73; see Section 2.4.5). Like the *surveillance* collocates in Figure 4-10, the words in Scott and Tribble's (2006) "KW linkage network" do not necessarily collocate with each other in a narrow span, but form meaning relationships nevertheless. The surveillance collocate network shares this more abstract representation of meaning relationships and emphasises the distribution of meaning across the corpus. The KKW groups in Section 4.3.2 achieved a similar high-level view of meaning components across the full corpus. The analysis of the shared collocates and their network has provided more knowledge of how meaning links develop across the S&S volumes.

The idea that the relationships between keywords and their distribution across the texts in a corpus is relevant to the text "prototypicality" approach put forward by Anthony and Baker (2015b) in the discussion of their tool *ProtAnt* (see Section 2.4.5). In relation to KW linkage networks, this prototypicality approach seems to take a different perspective on the same concept. If the concept of co-keyness emphasises the overall patterns, prototypicality helps to identify individual texts at the core of these patterns.

Brezina et al. (2015, p. 165) suggest that collocation networks are useful for both specialised and general corpora. The other existing methods work particularly well for specialised corpora (as argued for topic modelling by Murakami et al., 2017) or at least corpora of a specific genre (Scott & Tribble, 2006, p. 82). The *surveillance* collocate network developed in this section similarly reflects the lexical relations of a specialised corpus surrounding a KKW of a special status – *surveillance* – as even its salient collocates link to each other across

multiple volumes. This quality is quite distinct from the other S&S KKWs. Compared to the 19 shared collocates found for surveillance, only three are found for the second most frequent KKW, *social: exclusion, media* and *work*. However, these three collocates do not form a network, because they do not share their salience in any volume.¹⁴ The third most frequent KKW, *data*, does form a small network, which forms a triangle, as shown in Figure 4-11. These collocates relate to actions done to the data (*processing, protection*) and the type of data (*personal*). As with Figure 4-10, collocates that cross the boundary between Part I and II are shown in pink. *Personal* is only salient in the two first volumes and therefore shown in yellow. The triangle in Figure 4-11 differs from P. Baker's (2016) triangles (computed with *GraphColl*), in that his network connects words that collocate with each other in a given span. As a result, his triangles tend to contain "two words from the same grammatical class and/or having two of the words occurring in a lexical bundle or frame" (P. Baker, 2016, p. 155). By contrast, Figure 4-11 represents a more general network of the meaning constituents of *data* in the S&S Corpus.

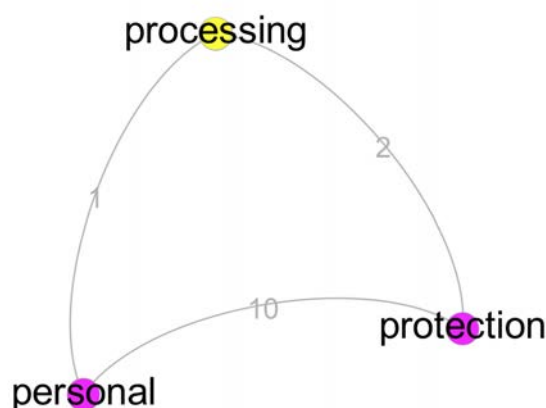


Figure 4-11: Shared collocates of *data*

¹⁴ *Exclusion* is a significantly more frequent collocate of *social* in Volumes 2 and 4, *work* in Volumes 5 and 7, and *media* in Volumes 9 and 12.

In summary, this section has introduced a network view of salient co-occurrence patterns across the S&S Corpus. As an approach based on evaluating co-occurrence patterns, it is methodologically similar to collocation networks and topic modelling. At the same time, the network of shared collocates resembles the abstraction that the KW linkage networks involve by design. I argue that this abstraction away from individual surface co-occurrence spans and examples relates to the quality that Murakami et al. (2017) describe as the “thematic meaning”. These attempts at exploring meaning apparently relate to Teubert’s (2019, p. 153) notion of ‘paraphrastic content’ for which, he suggests, we may have to “go beyond [...] the traditional notion of lexical meaning”. Perhaps the shared co-occurrence networks present one way of attempting that.

4.6 Conclusions

This chapter has analysed the representation of surveillance in the first discourse domain, academic discourse. As a corpus of experts’ accounts, the S&S Corpus has provided a useful overview of the concept of surveillance. In response to RQ 1-1, the chapter has shown that surveillance is a complex concept that can take many forms. The analysis of explicit definitions has identified intertextual references to broad definitions, highlighting the particularly popular definition of surveillance being “the focused, systematic and routine attention to personal details for purposes of influence, management, protection or direction” (Lyon, 2007, p. 14). The discussion about definitions involved questions such as how closely surveillance is related to watching and whether it is refined to human subjects. More specialised definitions are created for particular theoretical frameworks or practical applications of surveillance (e.g. in the public health context). As a result of this complexity, research in surveillance studies is portrayed as interdisciplinary. The chapter has identified KKWs as words that are consistently salient across

all volumes of the corpus in order to address RQ 1-2. These KKWs broadly relate to the description of (i) theoretical frameworks of surveillance, (ii) the government and public domain, (iii) the coordinates of time and space, (iv) academic discourse, (v) social actions and actors, (vi) techniques of monitoring and identification, (vii) technologies and (viii) notable theorists in surveillance studies. *Surveillance* and the ten next most frequent KKWs (*surveillance, social, data, security, CCTV, privacy, technologies, practices, individuals, cameras, space*) were found to cover many of these groups. As such, the KKWs provide a lexical approach to a high-level representation of the aboutness of the corpus and therefore the meaning of *surveillance* in this academic discourse. The KKW meaning groups relate to Mehrabov's (2015) themes, which are, however, more localised, given their origin in special issue titles from Volume 1–9. In terms of his themes, the co-occurrence patterns indicate a shift from “Classic Surveillance” at the beginning of the corpus towards more “Identity-based Surveillance” and the themes of resistance and empowerment in the later volumes including concepts like *participatory surveillance*. The theme of “Mobility and Stasis”, however, acts as a thematic link across the corpus. The analysis of the most recent volumes, not part of Mehrabov's (2015) study, suggests an additional theme of “Mass Surveillance” that would account for the discussion of surveillance after the Snowden “revelations” and developments of “big data” techniques, which are already foreshadowed by the KKW group of “technology”. The analysis of these shifting co-occurrence patterns has addressed RQ 1-3 and rounded off the chapter with a network perspective on the meaning components of surveillance in the S&S Corpus.

5 Places and mobility: Interaction orders in the discourse of surveillant landscapes of blog posts

5.1 Introduction

The concept of ‘place’ is crucial to all attempts to describe human experience. As Chapter 2 has argued, time and place are the two fundamental coordinates of social encounters. Both are central to meaning-making in discourse in general and surveillance discourse in particular. This is why place and time are each dedicated a chapter in this thesis. The present chapter focuses on spatial patterns and Chapter 6 is concerned with temporal patterns. Linguistically, lexicogrammatical resources for place and time are widespread. Even the nouns *place* and *time* themselves are highly frequent general nouns that appear in a variety of patterns with different functions (see Mahlberg, 2005). In surveillance discourse, spatiotemporal information is especially important, because it provides the key “coordinates” for logging any actions (see Lyon, 2007, p. 16).

This chapter makes a start on examining the discourses around these surveillance coordinates with a focus on place patterns. In a similar manner to McEnery’s (2006) combination of keyword analysis and moral panic theory (see Section 2.4.1), this chapter employs the theoretical framework of surveillant landscapes for guiding the linguistic analysis. Jones’s (2017) extension of the “linguistic” to a “surveillant” landscape emphasises that the physical environment not only provides input, but increasingly also “reads” any passers-by. As outlined in Section 2.6.1, the surveillant landscape framework entails three interrelated analytical perspectives, which are derived from mediated discourse analysis: (i) discourses in place, (ii) interaction orders, and (iii) historical bodies (see Jones, 2017, p. 153). The present

chapter proposes that the second perspective, the social, interactive dimension, is particularly crucial to the argument of the surveillant landscape as a concept. It emphasises the interaction between inhabitants and the landscape reading and writing each other.

This chapter analyses surveillance discourse with a focus on surveillant landscapes from two perspectives: (i) linguistic patterns surrounding a specific, physical place and (ii) the textual representation of surveillant landscapes more generally. For the first perspective, I focus on a local shopping centre in Birmingham, UK, the “Bullring & Grand Central”, because shopping centres are generally subject to high levels of surveillance. Focusing on a concrete place has allowed me to carry out a fieldwork visit to facilitate the mediated discourse analysis approach. Section 5.2 argues that there are parallels between mediated discourse analysis and corpus linguistics that make these approaches complementary. The section addresses RQ 2-1 (“How is the surveillant landscape multimodally represented in concrete examples?”) firstly with a multimodal analysis of the Bullring’s surveillant landscape according to the principles of mediated discourse analysis (5.2.1). In Subsection 5.5.2, I then demonstrate how corpus linguistic methods can provide additional insights into the textual representation of a particular surveillant landscape like that of the Bullring.

The second perspective moves away from the one particular place to consider the textual representation of surveillant landscapes on a larger scale, addressing RQ 2-2 (“How does the social dimension – the interaction order – contribute to the textual representation of surveillant landscapes in the Surveillance Blog Corpus compared to the S&S Corpus?”). The analysis focuses on the social dimension of surveillant landscapes, interaction orders. I view this dimension as central to the framework and consider it to be the most suitable focus for a corpus linguistic analysis of textual reactions to surveillant landscapes. Sections 5.3 and 5.4 present an analysis that is guided by the surveillant landscape framework but employs corpus linguistic

tools to apply these analytical categories to large amounts of textual data. This approach contributes to Barnard-Wills's (2012) calls for considering discourse and textual forms of linkages in the surveillant assemblage (see Section 2.2). The analysis focuses on surveillant landscapes in the Surveillance Blog Corpus (see Section 3.2.2) in comparison with the S&S Corpus of journal articles (studied in Chapter 4). The S&S Corpus has acted as a 'seed corpus' for compiling the Surveillance Blog Corpus. So, the blog posts have not been collected according to external criteria, but are based on internal, lexical criteria of surveillance discourse.

The comparison of blog posts with journal articles aims to contribute to the understanding of surveillance discourse in different public domains. Blog posts provide a useful source for the interactive features of surveillant landscapes, because they do not only link to many other sources (see Section 2.5.2; Myers, 2010), but are also published much faster than academic articles. So blogs can respond to events and stimuli in the social and physical environment more immediately, and potentially provide local coverage where the media fail to do so (see Section 2.5.2; Gordon & de Souza e Silva, 2011, p. 119). Yet, the virtual nature of blogs means that they are not restricted to writing from or about particular geographical locations. When place references occur, these are therefore likely to serve important functions, and they can point to virtual as well as physical places (Myers, 2010).

My analysis of surveillant landscapes in the blog posts in comparison to the journal articles builds onto the general co-occurrence methodology that was introduced in Chapter 3. This chapter extends this methodology from lexical to semantic co-occurrence comparisons. Section 5.3 introduces these semantic co-occurrence comparisons, which can be seen as a comparative approach to semantic tag collocation (see Section 2.4.6; Prentice et al., 2012). The new approach is explained in stages, moving from concrete to more abstract patterns. So, beginning with co-occurrence patterns of individual, concrete place nodes (5.3.1.1), the analysis

moves on to a manually grouped set of nodes (5.3.1.2) and a completely semantic comparison (5.3.2). The semantic stage is further divided into co-occurrence comparisons of specific tag subcategories (5.3.2.1) and, finally, the most general tags (5.3.2.2). Section 5.3, therefore, provides a thorough overview of patterns related to surveillant landscapes in the blog posts compared to the journal articles. Section 5.4 specifically examines the role of the social dimension of the surveillant landscapes and reflects on the corpus linguistic analysis of these interaction orders. Section 5.5 concludes the chapter.

5.2 Parallels between mediated discourse analysis and corpus linguistics

This section argues that the corpus linguistic approach to discourse analysis followed in this thesis has at least four parallels with the principles of mediated discourse analysis, which forms the basis for the surveillant landscape framework employed in this chapter. These parallels are summarised in Table 5-1.

Table 5-1: Parallels between mediated discourse analysis and a corpus linguistic approach to discourse analysis

	Mediated discourse analysis	Corpus linguistics
Discourse	language in use	
Unit of analysis	action	textual
Co-occurrence patterns	semiotic & material modes	textual
Register differences	affordances & constraints	linguistic comparison

First, for both mediated discourse analysis and corpus linguistics, a focus on “language in use” is crucial. In corpus linguistics, this principle is enacted through the discipline’s focus on empirical data collection and analysis. As pointed out in Chapter 2, a corpus linguistic approach views language as “a social phenomenon” (Mahlberg, 2005, p. 188). Similarly, mediated discourse analysis is characterised by a genuine interest in how people in the real world communicate and get things done in interaction with their environment. Importantly, this

research tradition actively moves beyond textual analysis by involving material and semiotic elements. This material focus is reflected in the terminology for mediated discourse analysis. So, analysts in this area prefer the term ‘sites of engagement’ over ‘context’ to describe the social, situational and physical surroundings, in order to move away from the implied association between ‘text’ and ‘context’ (Jones, 2012, p. 70).

Secondly, and related to this focus on language use, both approaches have a clearly defined unit of analysis, but do not analyse discourse exhaustively. Mediated discourse analysis focuses on the ‘mediated action’ (Jones, 2012, p. 28) such as crossing the street (Jones, 2012, p. 70) or ordering a meal, including all associated linguistic and material elements (e.g. the traffic light, the restaurant’s interior). The often-quoted coffee shop example by Scollon (2001, pp. 1–5) describes the action of buying a cup of coffee. Among other elements, the analysis considers the roles of various people (e.g. customer, barista, health inspector) related to the action as well as the role of language (e.g. the utterances used to order the drink and the conversations between friends, but also the messages printed onto the cups). The analysis cannot be exhaustive, but the definition of the unit of meaning – the action – is clear. By contrast, corpus linguistic analyses formally focus on textual units (e.g. words, sentences, paragraphs, whole texts). This textual focus is mainly based on convenience. In the past decade progress has been made with the development of multimodal corpora. These corpora make it possible to not only use the “the text as a ‘point of entry’ for corpus research” (Knight, Adolphs, Tennent, & Carter, 2008, p. 2), but also other modes like gesture. However, multimodal corpora are time-consuming to develop and are therefore restricted in scale (though see Bednarek & Caple, 2017 for new multimodal approaches to analysing news discourse).

In general, neither mediated discourse analysis nor corpus linguistics can analyse the entire discourse at large (see Teubert, 2010), because it is never possible to capture all

communicative situations. Moreover, it is unlikely that either method can analyse a single situation exhaustively. So, this chapter argues that the methods can complement each other to gain a fuller picture.

Thirdly, I propose that both approaches deal with different types of co-occurrence. One important concept in corpus linguistics is collocation (see Section 2.4.4), which is principally based on textual co-occurrence. This chapter argues that mediated discourse analysis can also be said to deal with co-occurrence: what material and linguistic elements occur together and in which constellation?

A fourth parallel is that both approaches are concerned with register differences. Due to the different units of analysis, a corpus linguistic approach focuses on the textual differences between registers by comparing linguistic features across texts (see e.g. Biber & Conrad, 2009). Mediated discourse analysis is concerned with the ‘affordances’ and ‘constraints’ of ‘cultural tools’ such as language, technology, etc. that affect what kind of actions somebody can take or how they communicate meaning in social situations (Jones, 2012, p. 68). So, a conversation via text message has the affordance of a more immediate response than exchanging letters, but is constrained in terms of nonverbal responses that are possible with face-to-face conversations.

The parallels between the approaches of mediated discourse analysis and corpus linguistics for analysing surveillant landscapes are best illustrated with an example. The following two subsections demonstrate the analysis of a surveillant landscape in a shopping centre, each with a different perspective. Section 5.2.1 follows Jones’s (2017) outline of the surveillant landscape framework based on the principles of mediated discourse analysis and 5.2.2 illustrates a corpus linguistic approach. The shopping centre is chosen as the focus of the case study because it is a (semi-)public place that has raised the attention of sociologists in general and surveillance studies scholars in particular. As Slater (1998, p. 149) argues, a

shopping centre “allows private finance to appropriate public space and to police it”. One of the measures taken for this “policing” of shopping centres is the use of CCTV, as research in surveillance studies has shown (see Section 2.2; Koskela, 2000; Walby, 2005). Signs announcing CCTV cameras are also found in my data along with other features of the surveillant landscape, as I explain in 5.2.1.

5.2.1 A mediated discourse analysis of the surveillant landscape in the Bullring shopping centre

As a qualitative illustration of the principles of mediated discourse analysis, this section focuses on the surveillant landscape of a particular shopping centre, the combined “Bullring & Grand Central” complex in the city centre of Birmingham. The Bullring is chosen as the largest local shopping centre and one of the largest in the UK as a whole (according to West Midlands Growth Company, n.d.). The following subsections illustrate what kind of data – collected from a visit to the shopping centre and its digital platforms – is relevant to an analysis of the discourses in place (5.2.1.1), interaction orders (5.2.1.2) and historical bodies (5.2.1.3) of this surveillant landscape .

5.2.1.1 Discourses in place

As the “semiotic dimension”, discourses in place cover any messages that the landscape communicates to passers-by (Jones, 2017, p. 154). Such messages may include alerts and warnings that surveillance is taking place, but also prompts for visitors to become “more compliant objects of surveillance” (Jones, 2017, p. 154). The Bullring is a large shopping centre

that was opened in 2003 in an area of Birmingham called the “Bull Ring”,¹ which has been hosting markets and other commercial activity since medieval times (“Bull Ring, Birmingham,” 2018). In 2016, the Bullring was connected to a smaller, newly developed shopping centre in the main train station (Birmingham New Street), called Grand Central. Although differences in design are found across Grand Central and the two wings of the Bullring, overall, the whole shopping complex is characterised by wide spaces and natural light. These corridors can be considered ‘passage spaces’ (Scollon & Scollon, 2003, p. 170), as unlike the space inside shops and cafes, people are meant to use the corridors to move along rather than to use them for specific purposes like shopping or eating. Therefore, there are few places to sit in the Bullring & Grand Central complex without being compelled to buy food or drink.

Figure 5-1 provides photos from my Bullring fieldtrip. The top left panel in Figure 5-1 shows a view of crowds on various floors of the one wing of the Bullring (“West Mall”) on a Saturday in July 2018. Accordingly, the general discourses in place include the typical elements of commercial linguistic landscapes such as shop fronts and advertisements (see Section 2.6.1; e.g. Lou, 2014). Examples include the large Debenhams panels featuring models in summer wear at the top of the photo and the digital advertising screen on the ground floor. The remaining panels of Figure 5-1 illustrate additional surveillant discourses in place that can be identified in the Bullring. Panel 2 shows an anti-terrorism message that is periodically displayed on the screen of the customer service counter asking visitors to “act on [their] instincts and call the police” if they “see or hear something that could be terrorist related”. So, this panel actively asks visitors to partake in surveillant activities. Panel 3 features a sign that announces the Bullring’s CCTV scheme, placed under a surveillance camera.

¹ The geographical area is spelled as two words, “Bull Ring”, but the shopping centre is spelled as one, “Bullring” (see “Bull Ring, Birmingham,” 2018).

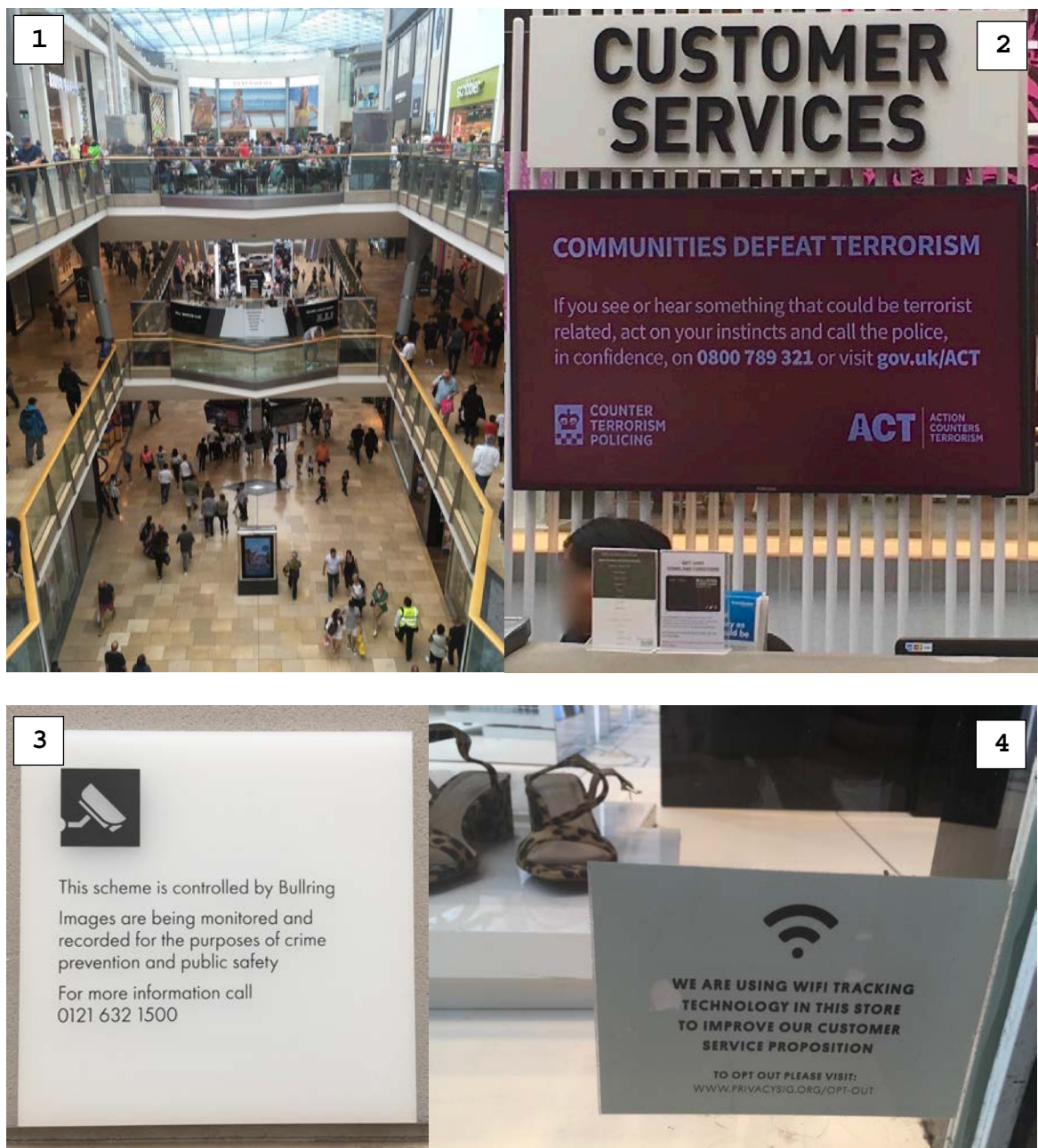


Figure 5-1: Photos from the inside of the Bullring and surveillance signs (taken on 28 July 2018)

Finally, Panel 4 shows a sign next to the entrance door to the clothing store Topshop stating that “WiFi tracking technology” is used in the store. Whereas the surveillance signs in Panels 2 and 3 appear in the general, open space of the Bullring, Panel 4 is an example of a store-specific sign. Many stores use their own security and surveillance technologies. Often, signs

can be found that announce the operation of CCTV. Yet, the WiFi tracking employed by Topshop seems to transcend the prevalent security purpose of surveillance techniques, as suggested by the statement “to improve our customer service proposition” (Figure 5-1, Panel 4). According to the brand’s website, visitors who are logged into their Topshop account on their mobile phone and/or have one of the company’s apps consent to Topshop “tracking your journey through our store by linking your journey and your account” (Topshop, n.d.). This sign can therefore be seen as one of the prompts for making oneself a more complacent surveillance object (see Jones, 2017, p. 154). The ‘emplacement’ of surveillance signs contributes to their indexical meaning in the environment (Jones, 2017, p. 164). Several signs in the Bullring, like the Topshop example, are emplaced in an inconspicuous way. Although Topshop provides customers the chance to “opt out” from the WiFi tracking, this part of the message is unlikely to be read. It is printed in a small font on a sign that may not be noticed in the first place, because the sign is placed very close to the floor.

In addition to the closed off individual shops, the shopping centre contains open shops in the main corridors that sell, for example, snacks or beauty products and whose shop assistants may stop passers-by to encourage them to try their products (such as a hair straightener at a stand in the Grand Central, see Figure 5-2 in the next section). These human actors contribute to the surveillant landscape. The following section explores, among others, examples of how people with surveillant roles affect the interaction orders of surveillant landscapes.

5.2.1.2 Interaction orders

Following the original mediated discourse analysis tradition, the ‘interaction order’ refers to the relationships between people in a social situation (see e.g. Jones, 2012, p. 70; Scollon & Scollon, 2004, p. 13; see Section 2.6.1). For example, the Bullring shopping centre is

particularly crowded on weekends – a situation which is likely to construct interaction orders that involve consideration (or annoyance) as shoppers have to make way for other visitors and their belongings (prams, shopping bags, etc.). Within the surveillant landscape framework, the concept of interaction orders extends to any type of relationship set up “between the watcher and the watched” (Jones, 2017, p. 169), which may therefore also be facilitated by architectural and technological features. The analysis then focuses on how these environmental features

[...] help to enforce particular power relations and particular sets of rights and responsibilities, and how they position different people as, for example, innocent, suspicious, desirable, or undesirable. (Jones, 2017, pp. 169–170)

Additionally, people can become part of these surveillant relationships when surveillance is carried out by “embodied actions of human beings” (Jones, 2017, p. 154). The shop assistants, in a way, act as human surveillant agents that look out for passers-by of interest – for example, people with long hair to target for promoting their hair straighteners. Figure 5-2 depicts this surveillant activity at a beauty stand with its shop assistants looking at passers-by. When the photo was taken another type of ‘embodied surveillance’ was coincidentally recorded as a pair of armed police officers were passing the stand (in addition to an unrelated passer-by on the right). The officers were on patrol through Grand Central, which is located on top of a large train station. As seen in Figure 5-1, anti-terrorism practices are, however, also present in the main Bullring shopping centre. Like for the WiFi tracking, a blend between commercial and security aspects is found in the surveillant landscape of the shopping centre.



Figure 5-2: The beauty stand in Grand Central (photo taken on 28 July 2018)²

The direct marketing approach of the beauty stand also illustrates one way in which the surveillant landscape can be face-threatening. Visitors may feel uncomfortable declining a sales offer, particularly if it is linked with a free service such as styling the hair. Another type of potential face threat in the surveillant landscape is linked with surveillance signs, like the Bullring CCTV sign in Figure 5-1. Such signs “must simultaneously address *both* the objects of surveillance and the beneficiaries of it, [...] they must construct their readers simultaneously as potential criminals and as potential victims” (Jones, 2017, p. 173; emphasis in original). The police officers similarly look at people with this double purpose. Accordingly, the passer-by is framed as a customer, a citizen protected by armed police as well as a potential terrorist that needs to be inspected.

² Ironically, my own act of taking this photo is also a form of embodied surveillance. I have added filters to anonymise all actors in this image.

This ‘multivoicedness’ (Jones, 2017, p. 173) is also reflected in the privacy policy on the Bullring website (“Bullring & Grand Central—Privacy Policy,” n.d.), particularly in the section pertaining to the processing of CCTV footage. Figure 5-3 illustrates how this section sets up various interaction orders with “visitors” who will be kept “safe and secure” on the one hand and “offenders”, whose “prosecution” will be assisted, on the other hand. The fact that the protected category includes “staff” and “property” illustrates that this policy serves the corporate interest at least as much as it protects the rights of the visitors. The protection of these groups and similar aims are framed as “legitimate interests” for processing the footage. This justification precedes and therefore appears to legitimise the announcement that the footage may be shared for the given reasons.

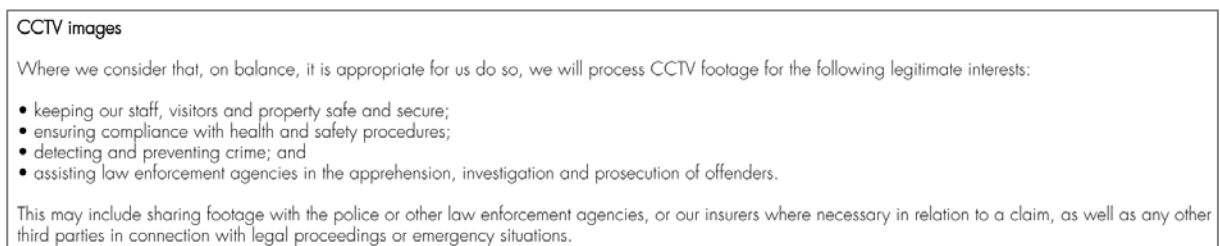


Figure 5-3: Extract from the Bullring Privacy Policy – “CCTV images”³

CCTV has been used in Britain for decades, initially mostly by retail before the government launched open street CCTV campaigns in connection with crime prevention in the 1990s (see Norris et al., 2004). More recently, though, “[m]obile phones and other digital technologies [have been] changing the nature of surveillant landscapes and the kinds of interaction orders they make possible” (Jones, 2017, p. 178). Like Jones’s (2017, p. 154) example of a bus company’s app that can collect passengers’ information beyond their journey, the surveillant

³ (“Bullring & Grand Central—Privacy Policy,” n.d.)

landscape of the Bullring involves an app. The “Bullring & Grand Central PLUS” app is a local version of an app rolled out across shopping centres in the UK and France (Godding, 2017). Given that the app features an “interactive map”, it is advertised strategically in the physical environment of the Bullring alongside a physical, static map on the centre’s signposts (see Figure 5-4).

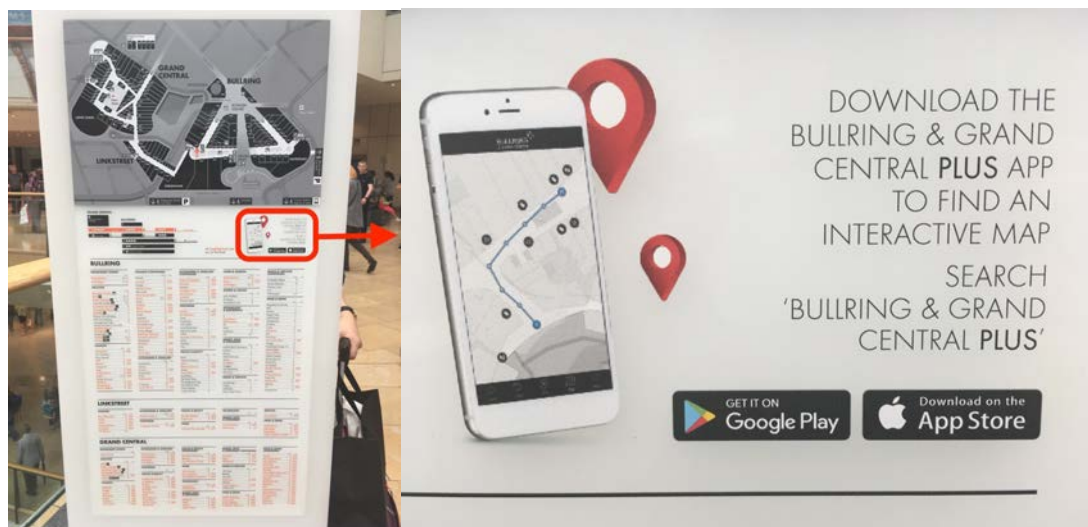


Figure 5-4: The physical signpost promoting the interactive map in the app (photos taken on 28 July 2018)

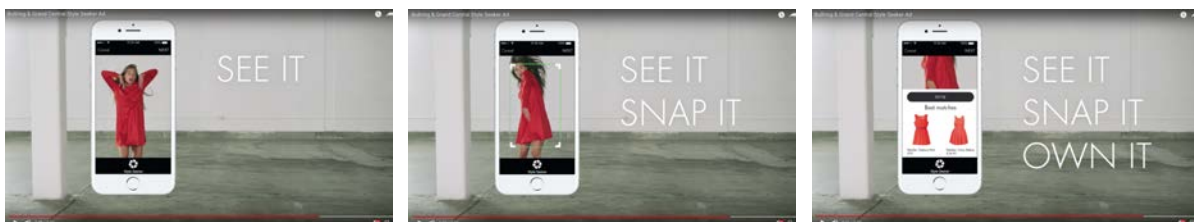


Figure 5-5: Still frames from the Bullring YouTube ad for the “Style Seeker tool” at 48, 49 and 51 seconds

One of the app’s features is called the “Style Seeker”. As advertised in a YouTube video (Bullring & Grand Central Birmingham, 2017), this feature allows users to take photos of outfits in order to receive shopping recommendations for a similar style in the Bullring. The description of the app refers to fashion items, implying that the user will only take photos of clothes, not of people: “take or upload a photo of an item you love” (App Store, 2017).

However, the promotional activities suggest that the app can be used to take photos of outfits worn by people, as for example illustrated in the still frames from the official YouTube video featured in Figure 5-5. Similarly, launch events of the Style Seeker feature in several shopping centres of the same chain with “living mannequins” (see e.g. Sixty9°, 2017; The Phoenix Newspaper, 2017) seem to have encouraged the practice of taking photos of people for checking their outfits in the app. Like Panel 2 of Figure 5-1 (Section 5.2.1.1), which asks visitors to report any suspicious behaviour, app users are encouraged to become potential watchers in the surveillant landscape with the Style Seeker. However, the interaction order that is created with this feature is more playful than that of the counter terrorism announcement, which places more responsibility on visitors.

5.2.1.3 Historical bodies

The effects that the surveillant landscape has on passers-by are captured by the third and final dimension of the framework, the historical bodies. Such effects can be psychological, i.e. how people internalise the surveillant architecture and its social relationships. These psychological effects are described as the ‘internalised’ historical body. In addition, there is an ‘externalised’ body of information building up about individuals as they pass through the surveillant landscape (see Jones, 2017, p. 180). Examples of both are evident in the Bullring data.

A particularly salient example of both internalised and externalised historical bodies is the shopping centre’s PLUS app. This app can provide convenient information for visitors of the Bullring, but also fulfils another purpose. A video on the Bullring website explains that the interactive map works by identifying users’ location via Bluetooth in relation to the centre’s physical space. Figure 5-6 shows two relevant still frames for which the voiceover is transcribed in Example (1).

- (1) Switch on Bluetooth and notifications every time you visit and PLUS becomes your real-time centre guide. It can even alert you to offers and promotions when you're near your favourite stores so you can discover savings plus other surprises and rewards, too.
("Bullring & Grand Central—Plus," n.d.; my transcription)

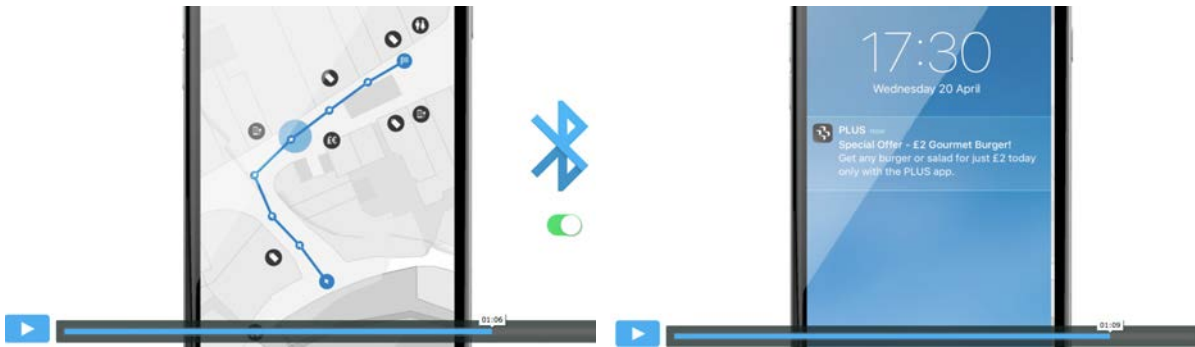


Figure 5-6: Still frames from the video “SEE EVERYTHING PLUS HAS TO OFFER” on the Bullring & Grand Central website⁴

In terms of the relationships constructed by the landscape, i.e. the interaction orders, the video constructs the user as a desirable customer. The user is rewarded with customised “offers” and “surprises”, such as the special offer for a gourmet burger in the notification shown in Figure 5-6 (right panel), when sharing locational information. This is not only relevant for the interaction orders of the surveillant landscapes, but also pertains to the aspect of the historical body. According to the shopping centre’s privacy policy, the app collects, among others, “Geolocation” data, as long as access permissions have been given via the phone’s settings. Figure 5-7 shows the section of the privacy policy that explains the ways in which the geolocation data is used, all of which seem to “inform and direct [the centre’s] marketing strategy”. The policy justifies the *use* of the gathered information with general benefits for the company (insights into “usage patterns”) and for customers (faster searches on the website, when they are recognised upon their return) as their digital historical bodies build up.

⁴ (“Bullring & Grand Central—Plus,” n.d.)

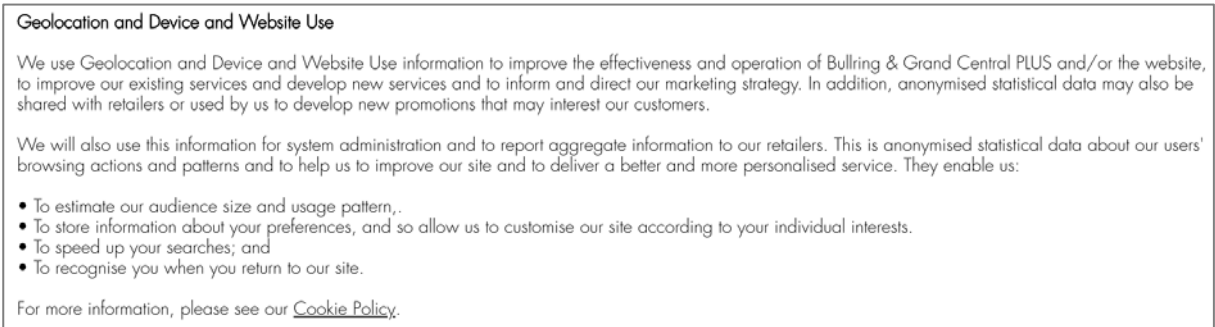


Figure 5-7: Extract from the Bullring Privacy Policy – “Geolocation and Device and Website Use”⁵

Accordingly, users of the PLUS app leave behind traces in the surveillant landscape of the Bullring. As seen with the Topshop WiFi tracking example in Section 5.2.1.1, these traces go beyond the CCTV footage that can capture all passers-by and extends to their own movements in the app if they are turning on the Bluetooth function. Users may build up an internalised historical body that encourages them to use the app again and return to the centre for customised offers, thereby affecting their future actions. Simultaneously, the externalised historical body of information about these users continues to grow. It is constituted of the trace of digital data that the operators of the shopping centre and its brands collect from their usage of mobile phones (via the app or WiFi tracking). So, the externalised historical body informs commercial strategies and can decide what future offers these particular users will be given.

Overall, Section 5.2.1 has shown that discourses in place of the Bullring’s surveillant landscape are found in a blend of physical and virtual spaces. Companies elaborate on their surveillant activities online, while only hinting at them on physical surveillant signs. Visitor data is (i) gathered for security as well as commercial purposes and (ii) via various digital and analogue channels. So, visitors are framed in multiple roles. The merging between the physical and virtual dimensions facilitates the building up of internalised and externalised historical

⁵ (“Bullring & Grand Central—Privacy Policy,” n.d.)

bodies. Companies benefit from the internalised effects that surveillant practices like the app can have on consumers. While customers may welcome the convenience of the app and its offers, the gathered data gives companies more power in the surveillant landscape, for example to improve their marketing strategies.

5.2.2 A corpus linguistic analysis of the surveillant landscape in the Bullring shopping centre

A corpus linguistic approach tends to be less focused on one particular place but considers mentions of a discourse object across a large number of texts. There can be exceptions, depending on the theoretical framework that the analysts work with. For example, Busse's (2019) work on the processes of place-making in Brooklyn, (see Chapter 2), is based on a specifically compiled corpus of interviews with residents in a particular neighbourhood. Equally, some studies in corpus stylistics majorly focus on a single text or a small set of texts such as a novella and a selection of its translations (see Section 2.4.2.1; Mastropierro, 2017).

Table 5-2: Top 15 most frequent word types in the Bullring privacy policy generated with *AntConc 3.4.4m*

Rank	Frequency	Type
1	109	to
2	87	you
3	74	and
4	70	the
5	70	your
6	68	we
7	67	or
8	60	our
9	57	of
10	49	data
11	34	any
12	33	for
13	33	personal
14	32	information
15	29	that

Accordingly, the Bullring privacy policy can be considered a mini corpus consisting of just one document with 2,549 words. Table 5-2 illustrates the “word counting” approach of corpus linguistics by listing the top 15 most frequent word types in this corpus generated with *AntConc 3.4.4m* (Anthony, 2016). Most of them are function words, as would be expected for any text. Yet, among these, second- and first-person pronouns (*you, your; we, our*) feature prominently. This suggests that the text attempts to create a personal interaction order by addressing the reader directly and using the third person plural pronouns instead of an impersonal reference to a legal entity. The discussion of this wordlist is a good example of how meaning emerges via comparison, according to meaning-making principle (ii) (see Section 2.7). The first- and second-person pronouns in Table 5-2 are so striking because they do not occur in the top ranks of every wordlist. For instance, the only pronouns in the wordlist of the S&S Corpus (see Table 4-1 in Section 4.2) are demonstrative and third-person pronouns.

The first content word in the list is *data*, followed by *personal*. These two words co-occur frequently as the compound *personal data*, as shown in Concordance 5-1. It is clear from this policy that historical bodies develop from the customers’ traces (see Section 5.2.1.3). The concordance lines illustrate that the policy (vaguely) discusses what happens to the data: it can be “retained”, “processed” and “shared”. Some of these instances are accompanied by face-saving hedges and modality, such as “We will *normally* retain your personal data” (line 26) and “We *may* share your personal data, *where necessary*” (line 28), so as not to alarm the readers. The policy further attempts to build up trust by emphasising that the operator is “committed to protecting” (line 25) the data and that “appropriate safeguards” (line 24) to facilitate this protection are in place.

<p>1 on your online account); • delete any 2 we:- • provide you with a copy of any 3 ral (we) are committed to protecting any 4 as providing appropriate safeguards for 5 ne-readable format; • have any incorrect 6 o are specifically contracted to process 7 ocuments referred to in it) explains the 8 onal data do we process We process the 9 ; where we no longer need to process the 10 s to process that information, where the 11 you do not want us to delete this 12 taprotectionofficer@hammerson.com. What 13 it is important for you to know what 14 articipate or otherwise provide us with 15 ach and every website that collects your 16 you with a copy of any of your 17 t; and • restrict the processing of your 18 nd/or • object to our processing of your 19 r request or with our processing of your 20 impact on you or our processing of your 21 tion imposed on us We will process your 22 be held We may store or process your 23 circumstances we will only process your 24 e appropriate safeguards to protect your 25 ough we are committed to protecting your 26 nformation We will normally retain your 27 al information with? We will share your 28 data on our behalf. We may share your 29 he acquisition. We will also share your 30 time. If we are going to use your 31 n in respect of the countries where your 32 rs or the wider public. Where will your</p>	<p>personal data that we no longer need for the personal data we hold about you, • provide you personal data you provide to us, including personal data, when we use service providers personal data corrected or completed (you can personal data on our behalf. We may share your personal data we process, what it is used for personal data set out below, provided to us by personal data, but you require the information personal data has been unlawfully processed or personal data; where we no longer need to personal data do we process We process the personal data we collect and how we use it. Personal Data in the above situations. You will Personal Data. You are solely responsible for personal data, provided by you, that we hold in personal data whilst we confirm whether any personal data, where we process it under the personal data, you are entitled to lodge a personal data, we will notify you. Any personal data to allow us to comply with any personal data outside the European Economic personal data for purposes for which you have personal data and ensure that it is compliant Personal Data, we cannot control and do not personal data is for a period of up to 7 personal data within the Hammerson group of personal data, where necessary, with personal data where: (a) you have provided Personal Data for any other additional purposes personal data is processed and/or copies of any personal data be held We may store or process</p>
---	--

Concordance 5-1: All concordance lines of *personal data* in the Bullring privacy policy

<p>22: in 3,000 franchised stores across 66 countries. 60: and delis in Birmingham Bullring opening hours 202: subsidiary Home What's On Bullring Birmingham The 44: colourful Australian tea shop has just opened at 27: and shoes. Debenhams - Lower Debenhams in the 29: not so high end prices. The modern store in the 46: round Thorntons - Middle The Thorntons store at 47: place for designer deals. The flagship store at 49: more than 300 Topshop stores in the UK and the 135: site was originally a Caffè Nero when the 20: Bullring Shopping and eating guide to Bullring 59: eat, park, shops Cafes and delis in Birmingham 136: has not only had to compete with its two other 150: Bullring and Grand Central - an area dubbed by 53: Browns Bar & Brasserie - Spiceal Street The 52: (Image: Iain Findlay) Read More Bill's in The 55: frozen yoghurt. Read More Mee-cha coffee shop 57: The Plain Jane Not Dog Read More Not Dogs in 2: after a man's body was found near Birmingham's 69: vertisement show ad British owner of Birmingham's 70: View comments The British owner of Birmingham's 71: Manchester. Popular: Hammerson owns Birmingham's 143: Nando's. Browns Restaurant can be found at The 158: B2 5JS Read More Shopping in and near Birmingham 50: price. Whistles - Upper The Whistles store at 14: and streets. We've put together this handy 36: modern, classic designs in its two floor 40: department selling perfumes and giftsets. The 41: worldwide. Click and collect is available in the 43: sport and art. There's a good selection at the 144: shopping centre We've put together this handy 32: March 2017. New Look Man will make that area of 145: for one hour. There's also a way to park free at 26: shoe and handbag store Coach is due to open at 54: famous 80s and 90s nostalgia cafe opened at 211: Building Society has closed its site next to the 16: Remember shops do come and go, so do check the 58: in Birmingham best places to park in Birmingham 56: like it. Read More Review: Mission Burrito at 45: dedicated soft play areas located just outside</p>	<p>Bullring is one of those stores, where the staff can Bullring is generally open from 10am to 8pm Monday to Bullring is a retail complex in Birmingham city centre Bullring, offering a choice of more than 200 teas and a Bullring offers wedding services, personal shopping, Bullring offers the best from Kurt Geiger's own collection Bullring offers boxed chocolates, personalised gifts, ice Bullring offers a large range of men's and women's Bullring one is a big one. It stocks everything from Bullring opened on September 4, 2003. In recent years, Bullring opening times Parking at Bullring 36 cool shops Bullring opening hours Bullring is generally open from Bullring outlets, but also the arrival of Pret a Manger Bullring owners as a 'Birmingham's Carnaby Street'. Bullring restaurant is the first Browns to open in the Bullring review: Does the city really need another chain Bullring review: Is this the worst coffee in Birmingham? Bullring review: It's food, but not as we know it St. Bullring shopping centre. The homeless man was found on Bullring shopping centre and London's Brent Cross fights Bullring shopping centre and London's Brent Cross will Bullring shopping centre (pictured) and London's Brent Bullring shopping centre We've put together this handy Bullring - shops, where to eat, parking Primark megastore Bullring stocks both women's and men's collections of Bullring store guide so you can make the most of your Bullring store. Religion - Linkstreet A grunge, East Bullring store has a 3 Mobile department too. Read More Bullring store. Read More Owners of Lasan restaurant to Bullring store. T2 - Upper This colourful Australian tea Bullring store guide to give you a complete store list, Bullring the place to go for men's clothing, as Zara Man Bullring too and details of new shop launches at the Bullring towards the end of 2017 and will offer the Bullring towards the end of 2016. Perhaps not to Bullring until September Birmingham City Council Bullring website beforehand to avoid disappointment. Do Bullring - where to eat, park, shops Cafes and delis in Bullring - will it pinch our best burrito crown? Mount Bullring within Moor Street Car Park - one for toddlers</p>
--	--

Concordance 5-2: Sample of concordance lines for *Bullring* from *WebCorp* with page filter "Birmingham", within the past six months (accessed on 13 June 2018)

The privacy policy is just one example of a text that is relevant to this particular surveillant landscape. As for the qualitative approach in 5.2.1, other related texts could include any signage in the shopping centre itself and any utterances produced at the place, as long as this linguistic data has been transcribed. More often, however, as was shown in the overview in Chapter 2, corpora are compiled with readily available electronic textual data such as social media texts or newspaper articles. For example, the *WebCorp* tool (Research and Development Unit for English Studies, 2017) can be used to generate an ad hoc concordance of *Bullring* from online newspaper articles. Using *Birmingham* as the page filter, most concordance lines originate from the local newspaper, the *Birmingham Mail* (see Concordance 5-2). They tend to have a promotional function, or at least provide information about what the shopping centre “offers” (lines 27, 29, 46, 47). The instances also provide practical information such as opening hours (lines 20, 59) and reviews (lines 52, 55, 57). In other instances, the Bullring is mentioned as a landmark in order to index the location of nearby incidents (such as the discovery of a man’s body; line 2).

Examples that are relevant to the surveillant landscape can be found when restricting the search to websites that contain both *Bullring* and *security*. In June 2018, this search yields only three instances from an article about a Bullring security guard that was suspended after “clashing” with a pregnant shoplifter (Concordance 5-3). So, the concordance highlights another example of embodied surveillance (like the police officers in Figure 5-2), with a particular focus on the interaction order and the power relationship between the actors. The event is newsworthy because violent behaviour from a security guards is not considered appropriate in relation to a vulnerable party such as a pregnant woman, even when guards are usually expected to act forcefully.

1: a Reach plc subsidiary Home News Midlands News Bullring News New Look SUSPENDS security guard after
 2: after brawl with pregnant shoplifter outside Bullring New Look has told the Mail that they have
 3: footage emerged of the clash outside the Bullring branch of New Look which shows the undercover

Concordance 5-3: All instances of *Bullring* with page filter *security* from *WebCorp* (accessed on 20 June 2018)

So, this section has shown ways in which corpus methods help with the analysis of textual representations of the Bullring's surveillant landscape. Concordance lines of the privacy policy have illustrated how the management tries to reassure visitors that their data is secure and only shared "where necessary". The examples from the *Birmingham Mail* are restricted to the interest of the newspaper editors and therefore relate to promotion and the reporting of security-breaching incidents. At the same time, the *WebCorp* results show that, as an individual shopping centre, the Bullring does not receive major news coverage. Consequently, only a limited amount of electronic texts about the Bullring is available for corpus linguistic analysis. The following sections therefore move away from the narrow focus on a specific (semi-)public place to a wider scope with an analysis of the surveillant landscapes represented in the Surveillance Blog Corpus in comparison with the S&S Corpus. As emphasised throughout this thesis, meaning emerges through comparative analysis.

5.3 Comparing the co-occurrences of place references in blog posts and journal articles

In this section, the focus of the chapter moves away from a particular surveillant landscape in the Bullring shopping centre to meaning-making patterns in the textual representation of surveillant landscapes more generally. For this purpose, this section introduces the corpus linguistic approach to co-occurrence comparisons of place patterns in surveillant landscapes. The following sections all focus on the Surveillance Blog Corpus. As this thesis views meaning as a comparative concept, the analysis compares patterns in the blog posts with those in another specialised corpus of surveillance discourse, the S&S Corpus analysed in Chapter 4.

The present chapter stays within the main methodological framework of the thesis: identifying co-occurrence patterns and comparing these across corpora in order to analyse meaning in a discourse. However, the perspective on co-occurrence taken here differs from Chapter 4. Instead of counting co-occurrences between individual word pairs, words referring to the same semantic field are grouped together for the co-occurrence comparison. So, the Surveillance Blog Corpus and the S&S Corpus have been annotated semantically (see Section 3.2.4). The semantic grouping makes it possible to extract references to places and mobility which I consider as the entry point to my corpus linguistic analysis of surveillant landscapes.

Section 5.3.1 illustrates how a co-occurrence comparison of semantically grouped nodes works, by manually grouping place words in the first instance. So that section starts with simple lexical co-occurrence comparisons, as carried out in Chapter 4, of nodes that have been qualitatively chosen as general place references and words referring to more specific public places and mobility-related terms. The analysis then moves towards semantic comparisons in stages, beginning with lexical collocates of manually grouped place nodes. Section 5.3.2 formally introduces comparisons with semantic tags. This way, the analysis moves from more “concrete” lexical to more “abstract” semantic results to examine the surveillant landscape.

5.3.1 Comparing lexical co-occurrences of place nodes

In this section, I explain the idea of comparing co-occurrences of semantically grouped nodes for my corpus linguistic approach to the surveillant landscape framework. I begin by comparing co-occurrences of individual place nodes (5.3.1.1) and then carry out the comparison for a manually combined PLACE node group (5.3.1.2). These sections represent the initial stages towards the complete semantic tag comparison that is introduced in Section 5.3.2.

5.3.1.1 Co-occurrences of individual place nodes

A lexical approach to analysing the representation of surveillant landscapes requires a relevant set of nodes as a starting point. Figure 5-8 shows the results for the nodes *space* and *place* in the blog posts compared with the journal articles. The results for the journal (on the right side of the plot) relate to the general focus on *urban* and *public space* in the surveillance studies literature, as illustrated in Concordance 5-4 (also see Section 4.4.1). Although the results for the blog posts do not indicate such a clearly coherent theme, they feature stylistic markers such as pronouns: *our* co-occurs significantly more frequently with *space* (see Concordance 5-5) and *you* with *place* (Concordance 5-6) in the blog posts.

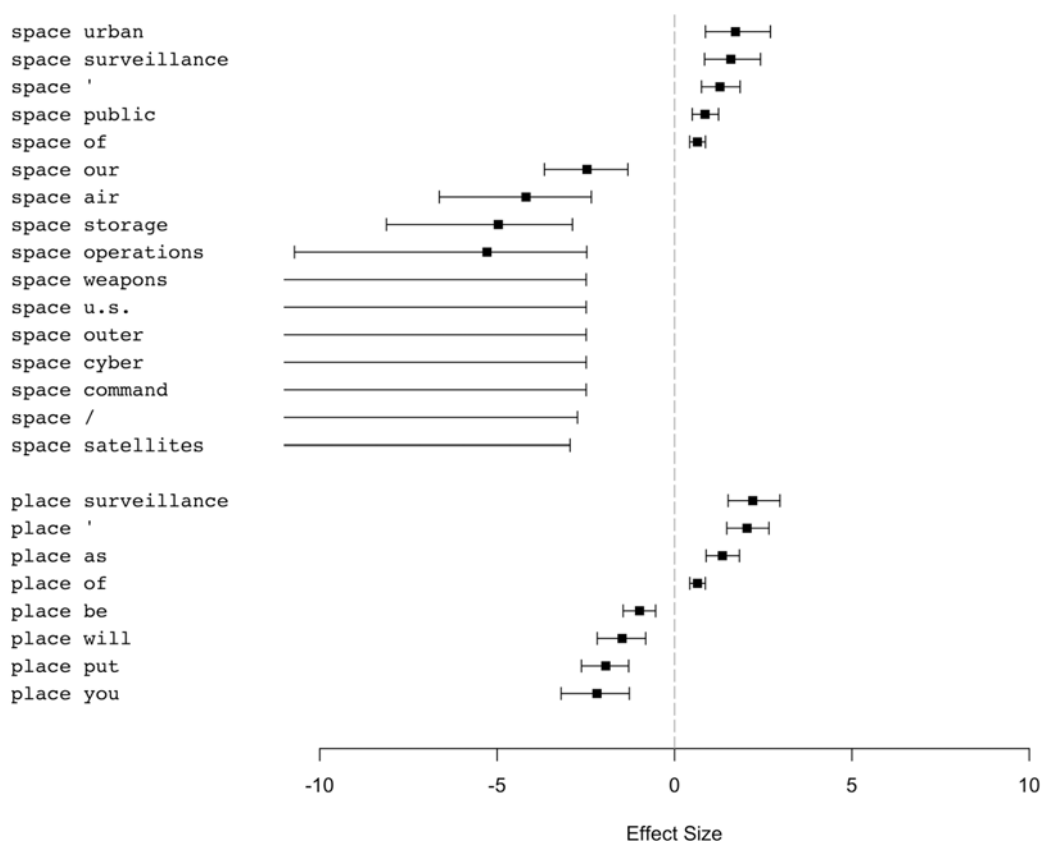


Figure 5-8: Lexical co-occurrence comparison of *space* and *place* in the Surveillance Blog Corpus (left) and the S&S Corpus (right)

[200] cameras in the (semi-) *public* space of the Zurich railway station
 [201] and activities from *public* *urban* space to private summer cottages, from
 [202]As *surveillance* spreads from material space to cyberspace the ' panoptic
 [203] any typical survey on *public* space *surveillance*. There was no doubt
 [204] role. In countries where *public* space *surveillance* is not forbidden (meaning

Concordance 5-4: *Space* co-occurring with *public*, *urban* or *surveillance* in the S&S Corpus (five of 732 examples)

[1] magnetic field extending out into space from *our* bodies, and that
 [2] magnetic field extending out into space from *our* bodies, and that
 [3] of a nuclear war. The Space Station which orbits *our* planet
 [4] invades the most precious private space we occupy: *our* homes. Utility
 [5] your DVR. How much storage space do I need ? *Our* handy
 [6] we think that the Internet space is privately *our* own because
 [7] the host of the Internet space can monitor *our* ' private
 [8] monitor *our* ' private ' space without *our* knowledge. Wednesday, October
 [9] *Our* model is efficient in space and can be stored in
 [10] framework and solutions in this space. *Our* analysis leads to several
 [11] foreign corps of *our* personal space in the non Euclidian space-temps
 [12] *our* experiences and conceptions of space, of embodiment (literally: of having
 [13] and *our* future energy economy, space exploration, and a host of
 [14] as, " a new kind of space, invisible to *our* direct senses,
 [15] to *our* direct senses, a space which might become more important
 [16] 10,000 Muslims in jail sharing space with terrorism convicts. *Our* foreign

Concordance 5-5: All instances of *space* co-occurring with *our* in the Surveillance Blog Corpus

[29] *you* comes at the right place. *you* can get special discount
 [30] to *you*. What About Work Place Privacy Laws ? Those generally are
 [31] activity. *You* ca n't legally place cameras in bathrooms or locker
 [32] can not go to your place of work, that *you* can
 [33] which, *you* 'll remember, took place in the months leading up
 [34] A WATCHFUL EYE--TEXTING, LIVE-STREAMING--ON YOUR PLACE WHILE *YOU* 'RE OUT. Canary
 [35] have a clear strategy in place that will help *you* avoid
 [36] about *you* in the first place ? And, more importantly, should it
 [37] it 's in the right place. QUESTION: Do *you* acknowledge
 [38] like *you* working in a place like that ? " Former Lockheed-NSA
 [39] do to install one is place it in the room *you*

Concordance 5-6: *Place* co-occurring with *you* in the blog posts (ten of 39 examples)

The general noun *place* can be used in a variety of patterns with different meanings (see Mahlberg, 2005, pp. 54–55, 143). It is used in phrases like *took place* (Concordance 5-6, line 33) and *in the first place* (Concordance 5-6, line 36), where *place* does not act as a noun referring to a particular location. *Place* is also used as a verb as in *place cameras in bathrooms* (Concordance 5-6, line 31). Similarly, *space* can occur in contexts that do not hold a strictly locational meaning, such as the compound noun *storage space* (occurring 14 times in the Surveillance Blog Corpus), illustrated in Example (2). The technological context of this noun phrase relates to the type of the texts in the corpus. Blog posts may be more often concerned

with the practical details of digital technologies than the more theoretically-focused journal articles.

- (2) Additionally, the massive amounts of *storage* **space** leave enough room that files may not even need to be erased [...]

(blogspot_outfile075; Vanacore, 2010)

Some issues of polysemy can be avoided by using more specific references to locations in the real world, e.g. related to public places and infrastructures. Figure 5-9 shows the results of a co-occurrence comparison for a more specific set of nodes (*airport, border, center, centre, city, district, mall, shopping, station, store, street, town* – as well as *park* and *shop*, which did not yield any significant results). These terms were chosen, because inner-city venues, shopping centres and transportation hubs regularly feature in the surveillance studies literature as heavily surveilled spaces (see the discussion of mobility surveillance in Section 4.4.1 and the Bullring case study in Section 5.2).

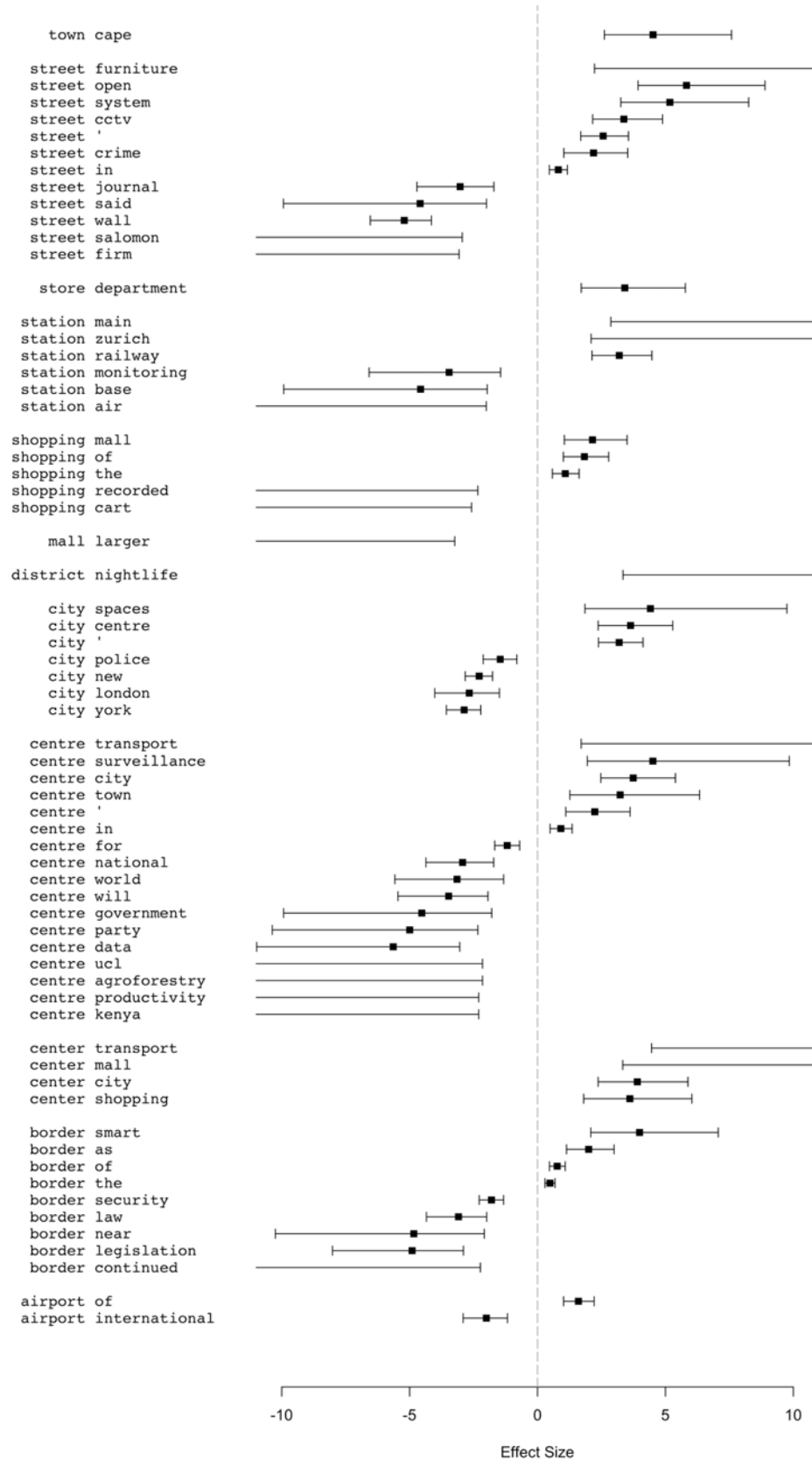


Figure 5-9: Co-occurrence comparison of lexical place nodes

Figure 5-9 indicates that several significant differences are found for most nodes. The node *centre* gives the most results. Some of these arise from recurrent proper names of institutions (e.g. *UCL Centre for Medical Image Computing*, *Centre Party*) and therefore demonstrate that the corpus patterns reflect social structures. In the journal, *surveillance* and *CCTV* are significantly more frequent collocates of *centre* and *street*, respectively. The blog patterns show other links with surveillance discourse: for example, the pair (*station*, *monitoring*) points to discussions of surveillance infrastructure such as “the principal NSA communications monitoring station” (wordpress_outfile155; PeterBDunn, 2014). One reason that the pair is more prominent in the blogs than in the journal is that some examples of *monitoring stations* have been copied from the Wikipedia page for “surveillance”. As the affordances of digital media make intertextuality easy (see Section 2.5.2), many blog posts contain quoted chunks from other sources. A different co-occurrence pair with an obvious relevance to surveillant landscapes in the blog posts is (*border*, *security*). Concordance lines of this pair indicate a political dimension to the discussion. Due to this political nature, the handling of border security can become the subject of criticism. This is illustrated in Example (3), which accuses a US government administration of lying about border security.

- (3) In July 2014, the Obama administration said:
 “Today **border** *security* is stronger than it ever has been.”
 Many of the entries on this list prove that the Obama administration was lying.
 (blogspot_outfile410; Warnings for America’s Future, n.d.)

Using individual word forms as nodes, only a few results are found that relate to the social dimension of the surveillant landscape framework. So, the next section groups the nodes together to explore this issue further.

5.3.1.2 Co-occurrences of manually grouped place nodes

When multiple lexical types are combined to a grouped node, more significant results are visible than for the individual nodes in the previous section. Figure 5-10 provides these grouped results, simply achieved by replacing all instances of the place nodes from 5.3.1.1 (i.e. *airport*, *border*, *center*, etc.) with the group label PLACE in the two corpora. This step achieves a simple, “manual” form of semantic tagging to illustrate the mechanism (see Section 5.3.2 for the full semantic comparison).

So, the accumulated co-occurrences of the various nodes produce additional results. New collocates of the PLACE group that are not listed in Figure 5-9 at all include, for example, *apparatus*, *securitization*, *private* and *public* on the side of the journal, and *privacy*, *command*, *medical*, *force*, *wall*, *securities*, *FBI* and *Obama* on the side of the blog posts. Concordance 5-7 illustrates that a range of the PLACE types actually occur together with *privacy* in the Surveillance Blog Corpus: *airport*, *border*, *center*, *city*, *shopping*, *shop*, *store*, *street*. Therefore, the grouping of nodes allows more patterns to be identified in the comparison that would otherwise remain unnoticed.

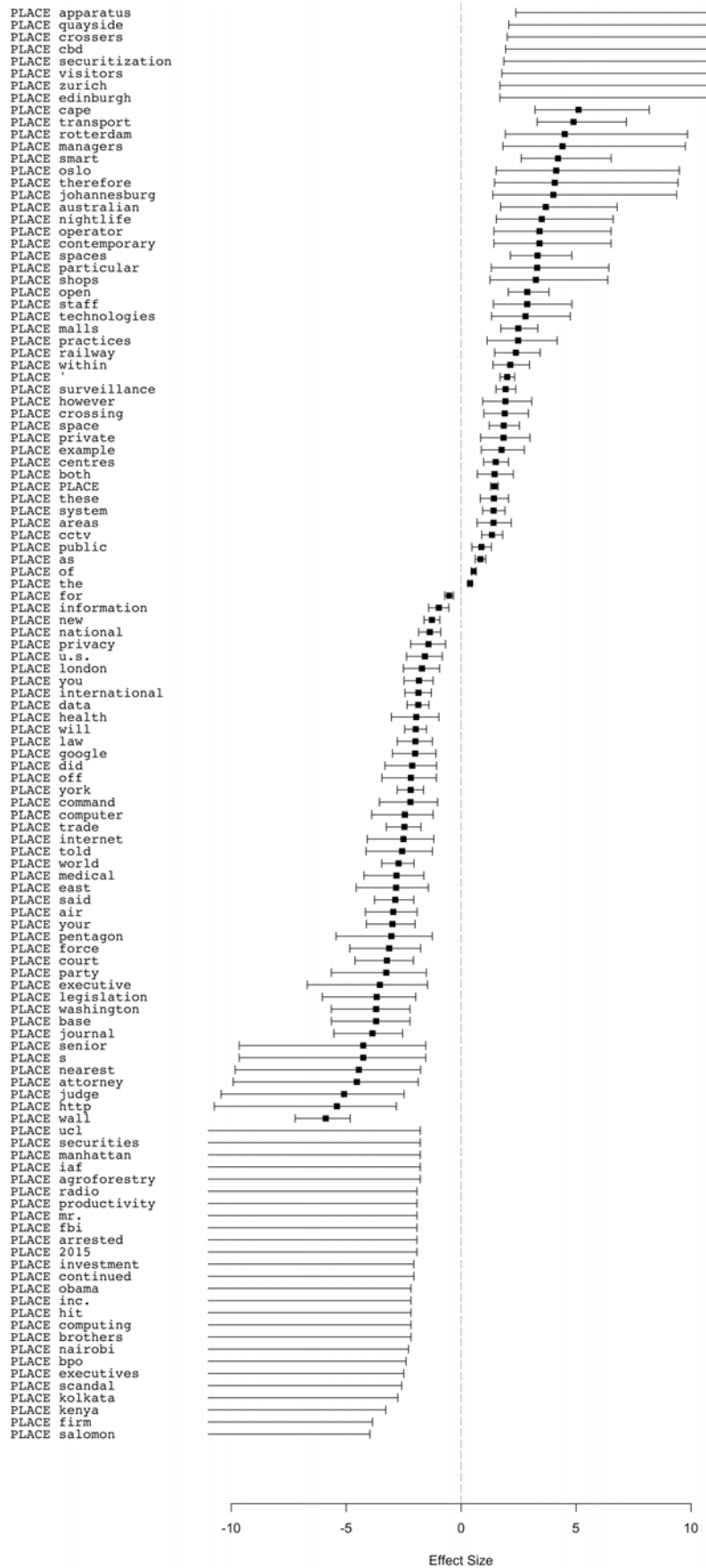


Figure 5-10: Co-occurrence differences for PLACE nodes (grouping the individual types from 5.3.1.1)

[1]	speaking loudly on a public	street , I have waived my *privacy*
[2]	as Alvaro Bedoya from the	Center on *Privacy* and Technology at
[3]	to the Electronic *Privacy* Information	Center , within this structure, " suspicious
[4]	policies and pages. Their *privacy*	center claims that it has 5
[5]	of the Electronic *Privacy* Information	Center (EPIC), and with help from
[6]	orts the Electronic *Privacy* Information	Center , a non-profit currently suing the
[7]	of the Electronic *Privacy* Information	Center (EPIC) articulates why reducing
[8]	ise, the Electronic *Privacy* Information	Center (EPIC) works to " focus public
[9]	ite. The Electronic *Privacy* Information	Center website. Available online: http:
[10]	be allowed to share Canadian	border into under new *privacy* charter
[11]	be allowed to share Canadian	border info under new *privacy* charter
[12]	News - U.S. Can share Canadian	border info under *privacy* deal - The
[13]	be allowed to share Canadian	border info under new *privacy* charter -
[14]	start. Google 's active *privacy*	center educates users about how it
[15]	rding to Electronic *Privacy* Information	Center 's Amie Stepanovich, " Currently, the
[16]	s. " The Electronic *Privacy* Information	Center (EPIC) obtained a partially redacted
[17]	profit Electronic *Privacy* Information	Center . " My concern is about discrimination,
[18]	this invasion of *privacy*. Main	Street Safety, an internet based company
[19]	have no *privacy*. New York	City ' s Central Park leads
[20]	A THOUSAND CUTS: HOME INTRUSIONS,	STREET THEATER, LOSS OF *PRIVACY*, VANDALISM,
[21]	*privacy* suffered by the owners.	Border Patrol and DHS has taken
[22]	planning to launch an online	store called PRCVM (short for " *privacy*
[23]	collaboration with the New Museum	store called " the *Privacy* gift shop. "
[24]	store called " the *Privacy* gift	shop . " Some other products from that
[25]	by the Electronic *Privacy* Information	Center and the American Civil Liberties
[26]	The Naked Ape in the	Airport Personal *privacy* is a thing
[27]	from the Electronic *Privacy* Information	Center - In May 2004, the Department
[28]	at the Electronic *Privacy* Information	Center .. Concerning Echelon 's inherent
[29]	Act, the Electronic *Privacy* Information	Center (EPIC) discovered that the agency
[30]	on-based Electronic *Privacy* Information	Center accused the FBI of not
[31]	tfit the Electronic *Privacy* Information	Centre (EPIC), under a Freedom of
[32]	sed. The Electronic *Privacy* Information	Center (EPIC), which sued the FBI
[33]	rded the Electronic *Privacy* Information	Center (EPIC) \$500,000 in settlement funds,
[34]	acy* guidelines, platform and application	store requirements, as well as relevant
[35]	Pulls Apps from Chrome Web	Store Over *Privacy* Issues Google has
[36]	*Privacy* Policies Misunderstood The Pew	Center for Internet and American Life
[37]	plan to do less online	shopping due to *privacy* concerns and
[38]	it surveillance friendly ? " [The Wall	Street Journal] See also: [*Privacy* groups
[39]	t protected, " said Georgetown University	Center on *Privacy* and Technology Executive
[40]	the report states. [The Wall	Street Journal] CA - CancerLinQ, *Privacy*
[41]	licy The Electronic *Privacy* Information	Center (EPIC) has filed a complaint
[42]	S.-based Electronic *Privacy* Information	Center (EPIC) submitted a complaint to
[43]	of the Electronic *Privacy* Information	Center said, " This is no time
[44]	Text The Electronic *Privacy* Information	Center (EPIC) has filed a complaint
[45]	Problematic for *Privacy* A Litigation	Center of the AMA and State
[46]	Japan causing Google to adapt	Street View to respect *privacy* concerns.

Concordance 5-7: All manually grouped PLACE nodes co-occurring with *privacy* in the blog posts

Importantly, the simple manual grouping approach used here relies solely on word forms. Although the group results provide more information than gained for individual nodes (Figure 5-9), Concordance 5-7 contains one example where this lexical approach is too simplistic: line 37 contains the node *shopping* not as part of the intended place compounds *shopping centre/mall* but as an activity. This particular case hints at the bigger issue that word forms and meaning do not form a “one-to-one relationship”, because word forms can be polysemous (Mahlberg, 2005, p. 189). In contrast to the “manual” grouping of a set of word

forms, automatic semantic tagging takes additional information into account (such as the word's part of speech and its context; see Section 2.4.6). So, a co-occurrence comparison based on semantic tags has the advantage that it only compares those uses of a word form which have a high likelihood of realising the intended meaning sense.

5.3.2 Co-occurrences of automatically tagged “M” nodes

The aim of the following sections is to identify co-occurrence patterns about surveillant landscapes on a larger scale. For this step, semantically grouped nodes are introduced and their co-occurrences are examined. In order to carry out a semantic co-occurrence comparison, the Surveillance Blog Corpus and the S&S Corpus were annotated with the *USAS* tagger (see Section 3.3.4). In order to assign a word to one of 21 major discourse fields (Archer et al., 2002, p. 2), the tagger makes use of information like the a lexicon in addition to the word's POS and context (see Section 2.4.6; Rayson, 2003). These major discourse fields are represented by capital letters in the semantic tags. More specific subcategories are indicated by decimal numbers. For some tags, plus and minus signs additionally locate the tagged word along a ‘semantic scale’ (Archer et al., 2002, p. 1), e.g. from “Helping” (“S8+”) to “Hindering” (“S8-”).

Table 5-3 illustrates the tagging with an extract from a book by surveillance studies scholars aimed at the public (Bennett, Haggerty, Lyon, & Steeves, 2014). The shaded rows contain the original text, with the semantic tags provided in blue for each token. All determiners and prepositions, for example, are assigned “Z5” as a “grammatical bin”, unsurprisingly forming the largest group, and pronouns are assigned “Z8” for “pronouns etc.”. The most relevant tags for this chapter include the “M” tags (“Movement, location, travel & transport”), shown in bold, which form the nodes for the analyses of the co-occurrence comparison as a proxy for mobility- and place-related references. In Table 5-3 only the subcategory “M6”

(“Location and direction”) is present (for the word forms *anywhere* and *where*). Because of the analytical focus on the interaction orders of the surveillant landscapes, this chapter pays particular attention to collocates from the “S” (“Social actions, states & processes”) and “E” (“Emotional actions, states & processes”) tags, to which the collocates are restricted in Section 5.4. Table 5-3 contains some bold instances of “S” tags, such as the verb *allow*, which represents a subcategory of the S7.4 “permission” tag. If a word is not recognised by the tagger it is assigned “Z99”, the “unmatched” category. This has only happened for one token in Table 5-3, *caf* in the final line; itself a “broken” token because the online interface did not deal with the *é* character in *café* correctly.

Table 5-3: Extract from Bennett et al. (2014, p. 87), semantically tagged with the *USAS* tagger⁶

Until	about	five	years	Ago	,	a
Z5	A13.4	T1.1.1[i1.3.1	T1.1.1[i1.3.2	T1.1.1[i1.3.3	PUNC	Z5
favourite	claim	of	Internet	Pundits	was	that
E2+++	Q2.2	Z5	Y2	X9.1+/S2mf	A3+	Z5
new	information	and	communication	technologies	(ICTS
T3-	X2.2+	Z5	Q2.1	Y1	PUNC	Z3c
)	would	make	geography	Irrelevant	.	They
PUNC	A7+	A1.1.1	P1/W3	A1.2-	PUNC	Z8mfn
envisioned	a	world	where	New	technology	would
X2.6+	Z5	W1	M6	T3-	Y1	A7+
allow	us	to	easily	communicate	with	people
S7.4+	Z8	Z5	A12+	Q2.1	Z5	S2mfc
anywhere	in	the	world	,	get	the
M6	Z5	Z5	W1	_PUNC	A9+	Z5
information	and	media	we	Desire	from	anywhere
X2.2+	Z5	Q4c	Z8	X7+	Z5	M6
,	and	work	equally	Well	in	the
_PUNC	Z5	I3.1	A6.1+	A5.1+	Z5	Z5
office	,	at	home	,	or	in
I2.1/H1c	PUNC	Z5	H4/H1c	PUNC	Z5	Z5
a	caf	in	Antigua	.		
Z5	Z99	Z5	Z2	PUNC		

My focus on semantic tag co-occurrences is similar to Prentice et al.’s (2012) concept of semantic tag collocation (see Section 2.4.6), but differs in two points. First, Prentice et al.

⁶ Using the Wmatrix (version 3) interface (Rayson, 2009).

(2012) focus on co-occurrence relations between a tag and a word. The present chapter analyses co-occurrences between semantic tags (as both nodes and collocates). Secondly, their approach is based on a traditional collocation association measure within one study corpus, but I compare semantic co-occurrences between corpora. In addition, there are differences in the thematic focus. Although both approaches are concerned with places, Prentice et al. (2012, p. 274) propose a method that can aid counter-terrorism research by identifying places that are at risk of attacks. By contrast, this chapter is concerned with the role of place patterns in surveillance discourse.

The following sub-sections present the results of the semantic co-occurrence comparisons. The first subsection (5.3.2.1) examines co-occurrence comparisons of two subcategories of the “M” tag that are relevant to the representation of the surveillant landscape, because they depict places and air travel. Subsection 5.3.2.2 then presents the most abstract co-occurrence comparison at the level of the major discourse fields in the *USAS* tagset (e.g. “A” and “M”).

5.3.2.1 Co-occurrences of semantic tags for places and air transport

This section replicates the comparison of manually grouped PLACE nodes in Section 5.3.1 on a larger scale based on automatic semantic tags. In order to study the representation of surveillant landscapes in the blog posts, I have chosen the “M” tag (“Movement, location, travel & transport”) as the tag that is apparently most relevant for mobility and public places. For the purposes of this chapter, the “Z2” (“Geographical names”) nodes seem too geographic and general to be useful as pointers for surveillant landscapes. As Table 5-4 illustrates, the “Z2” tokens in the Surveillance Blog Corpus are strongly associated with particular country and city names. While these may be useful nodes for purposes like understanding extremist threats for

particular places (see Prentice et al., 2012), nodes related to mobility and urban spaces are deemed more useful for this chapter.

Table 5-4: 20 tokens of the tag “Z2” (“Geographical names”) from the beginning and end of the Surveillance Blog Corpus, respectively

First 20 Z2 tokens	Minneapolis, Ringwood, UK, London, Canary, Minnow, American, Asia, Australia, England, Zion, America, Citadel, Utah, Canada, Asian, Russia, China, N.H., Israeli
Final 20 Z2 tokens	Labrador, Longley, Latvian, Hoes, Cuba, Playa, Tyneside, Bakersfield, Nauru, Liberian, Monrovia, Karachi, Ulster, PISA, Scandinavian, Nordics, Anfield, Bootle, Netherton, Southport

The semantic tag “M” refers to various broad concepts relating to mobility. Some of its subcategories include high frequency words such as *where* and *through* (both “M6”). In order to identify more specific meaningful patterns, it is useful to start the analysis with the individual subcategories. These are listed in Table 5-5 together with examples of the tokens by which they are realised in the blog posts. The examples illustrate that many of the subcategories are more concerned with movement or direction than places per se. So, I will focus on two subcategories in particular that are most relevant for surveillant landscapes. First, I analyse co-occurrences of “M7” (“Places”), because the focus on place is a good starting point to investigate the patterns of surveillant landscapes. So, the “M7” tag in particular can be considered a larger-scale, automatic version of the manually grouped PLACE tags in Section 5.3.1. Secondly, I examine co-occurrences of subcategory “M5” (“Movement/Transportation: Air”). This tag is a similarly relevant category for surveillance discourse, because airports are often discussed in relation to security checks and border monitoring (see Section 4.4.2.1).

Table 5-5: The first 20 distinct realisations of each “M” subcategory in Surveillance Blogs Corpus (ordered alphabetically)

Tag	20 examples from the blog posts
“M1” (“Moving, coming and going”)	access, advanced, come, coming, dispersion, expeditions, fall, go, leave, leaving, left, move, out, repatriation, returning, revisiting, through, up, walked, went
“M2” (“Putting, taking, pulling, pushing”)	brought, held, hold, holding, imported, laid, out, places, put, puts, raised, raises, remove, removed, send, sends, sent, shift, shifting, throwing
“M3” (“Movement/Transportation: Land”)	a, carrier, carriers, cars, cyclotron, drive, driven, drivers, drives, firewall, firewall., ⁷ passengers, path, pathways, platform, road, street, streets, traffic, train
“M4” (“Movement/Transportation: Water”)	brig, flow, frigate, grounded, grounds, helms, jetty, launch, launched, launching, naval, navigation, ports, sails, ship, stern, surf, surfing, vessel, wake
“M5” (“Movement/Transportation: Air”)	a, aerospace, airline, airlines, airport, baggage, basis, bomber, bombers, companies, fighters, flight, fly, handler, helicopters, on, out, pilot, pilots, terminals
“M6” (“Location and direction”)	above, anywhere, below, course, direction, elsewhere, everywhere, faced, ground, grounds, left, librarianship, misplaced, off, on, out, somewhere, stands, this, where
“M7” (“Places”)	aboriginal, area, boundaries, call, campus, cities, foreign, homeland, indigenous, islander, local, national, native, of, ports, site, sites, town, zone, zones
“M8” (“Remaining/stationary”)	abiding, deterministic, down, linger, lingering, loitering, parking, perched, sat, settle, settled, settles, sit, sits, sitting, static, stay, staying, stays, still

The results for the first tag I focus on, “M7”, are given in Figure 5-11. Because the “M7” comparison generates a lot of results, this plot has been reduced to only show co-occurrence pairs that are more than twice as different (i.e. an effect size over one). The largest difference on the side of the blog posts is the “A8+” tag, which relates to appearance and impression. Yet this is based on little evidence (see the wide confidence interval). The “G3-” (“against warfare”) shows the next biggest difference with more evidence. This tag is often realised by the collocate *civilian*. Example (4) illustrates the use of “G3-” in describing a surveillant landscape of civilian areas that are potentially at risk of police monitoring via drones. The account of the technical specifications of the police surveillance equipment together with the comment “privacy

⁷ *Firewall* appears twice here, because the *CLAWS* tagger has wrongly tokenised one instance as a “unit of measurement” (NNU) ending in a full stop like an abbreviation, rather than a common noun.

advocates worry” suggests concerns about the interaction orders in this surveillant landscape. The example negatively frames the possibility that the police may arrange surveillance measures which can “read” civilian residents from the air.

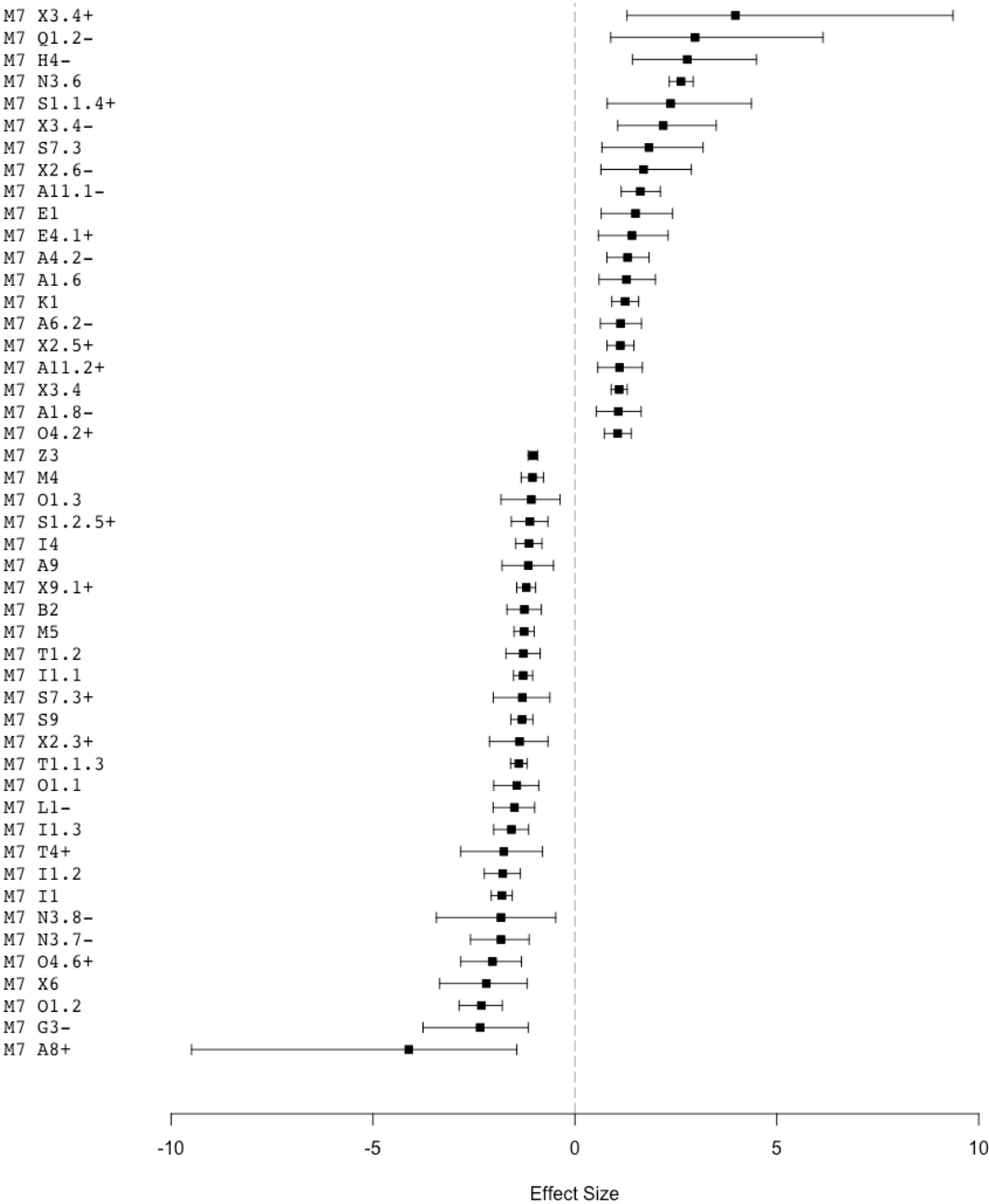


Figure 5-11: Significantly different co-occurrences of the “M7” (“Places”) nodes in the blog posts (left) and journal articles (right) with an effect size over one

- (4) [...] the Federal Aviation Administration (FAA) plans to implement new rules that would allow the routine flying of these drones across the United States by 2013; equipped with high-resolution, infrared and thermal-imaging cameras, these drones could provide police with the accurate monitoring of all types of **civilian* areas* and topographies; privacy advocates worry

(blogspot_outfile257; ASC, 2014)⁸

The “X6” (“deciding”) tag shows a similar salience in the blog posts. Most instances of “X6” co-occurring with the “M7” places are realised by forms of *ESTIMATE*. Example (5) illustrates one of these co-occurrences in the context of security operations by the police and the FBI on US college campuses. The example is written in a news article style, as it appears to have been reposted from a newspaper website. It points to embodied surveillance by campus police and FBI on college campuses as a response to the “Sept. 11 attacks”. The described “strengthening” of the relationships between these departments in the name of security reflects a reinforcement of the surveillant landscape.

- (5) But the Sept. 11 attacks prompted the bureau to strengthen its links to local and state police departments, including those on college campuses. Precise numbers are not available, but FBI **estimates** and interviews with **campus** police administrators indicate that at least a dozen departments have assigned officers to play significant roles in FBI anti-terrorism task forces.

(blogspot_outfile242; Roger, n.d.)

Among the salient semantic collocates in the blog posts are literal or metaphorical descriptions of material characteristics: e.g. “O1.2” (“Liquid”), “O4.6+” (“Temperature: Hot/ on fire”), “N3.7-” (“Length/Height”), “N3.8-” (“Speed: Slow”). Whereas the liquid tag tends to refer to oil or water in political and economic contexts, the other tags are used in more diverse circumstances. Example (6) illustrates a promotional use of an “N3.7-” tag in the description

⁸ For this and following examples of semantic tags, words in bold still represent nodes (here, *areas*) and collocates are shown in asterisks. So the “G3-” collocate in this example is “G3-” *civilian*.

of surveillance gadgets (infrared cameras). The tagged adjective *low* here refers to the adverse conditions in which other surveillance cameras may not function as desired, but for which this type is “well suited”. Like Example (5), Example (6) implies a desire for a stronger surveillant landscape, but represents a less institutional form of that.

- (6) These cameras are well suited for surveillance of *low* light **areas** or **areas** with no light at all.

(blogspot_outfile037; Bhadra, 2010)

Other salient semantic collocates of “M7” in the blog posts relate to money and business: “I1” (“Money generally”), “I1.1” (“Money: Affluence”) and “I1.2” (“Money: Debts”). Collocates belonging to the general “I1” include words like *financial*, *monetary*, *currency*, *accounts*, *dollars* and *commercial*. One context in which the (“M7”, “I1”) pair appears is financial crime as a subject of surveillance and investigation.

Figure 5-11 further contains subcategories of the “Social actions, states and processes” (“S”) tag relating to religion (“S9”), competition (“S7.3”) and strength (“S1.2.5”). These can be considered references to the social dimension, the interaction order, of surveillant landscapes (see Section 5.4). Strikingly, two forms of the same tag appear on different sides of the plot: the competition tag as a whole (“S7.3”) is significantly more frequent in the journal articles, while the subcategory, “S7.3+” (“Competitive”) is a more frequent “M7” collocate in the blog posts. Concordance lines suggest that the collocates indeed refer to different forms of competition. In the blog posts, the “M7” nodes co-occur with forms like *compete*, *competitive* and *competitiveness*. These tend to refer to economic competition (e.g. “a vibrant and globally competitive financial sector”). Conversely, the “S7.3” tag is mainly realised by forms of CONTEST in the journal articles, in the context of “contested urban sites” or “whether or not these borders are contested”. So, in the blog posts, there appears to be economic competition

between different places. In the journal, the overall competition tag relates to the negotiation of concepts and places.

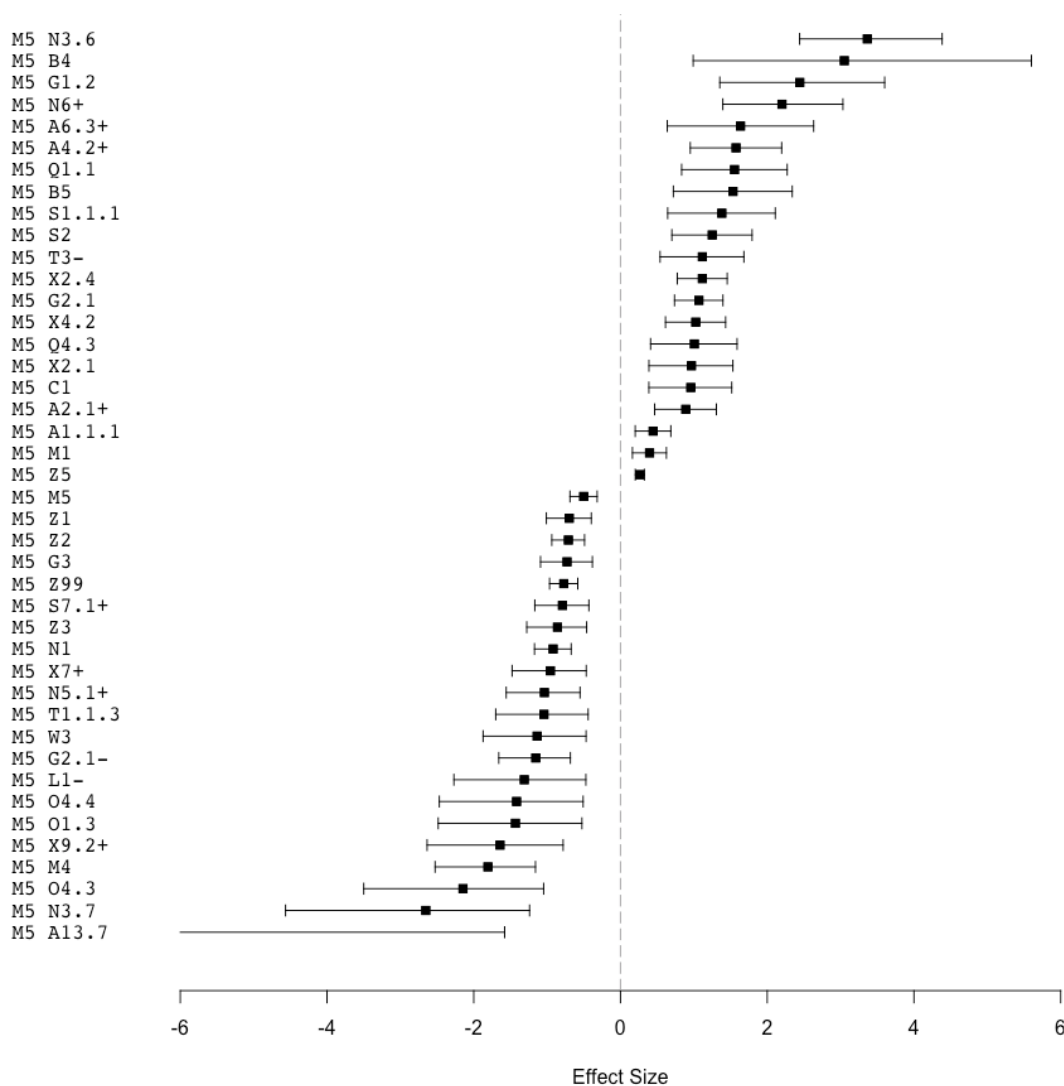


Figure 5-12: All significantly different co-occurrences of the “M5” (“Movement/transportation: air”) nodes in the blog posts (left) and journal articles (right)

The results for the other node that is particularly relevant for the analysis of surveillant landscapes, the “M5” tag on air travel, are presented in Figure 5-12. Among the most different collocates are references to colour (“O4.3”) and length/height (“N3.7”), which tend to appear in aviation descriptions (“a small **white** plane”; “**black** eagle flight”; “**green** jet fuel”;

“high *altitude* **flight** mode”). The tag “X9.2+” (“Success”) also belongs to this category, as it is mainly realised by the token *take off*, which is used literally rather than metaphorically here: e.g. “*take off* from airfields in other countries”. The difference for these categories is likely to be so large, because the journal articles have a narrower focus on the surveillant aspect of air traffic, rather than discussing plane types or flight routes.

The tag “X7+”, also a significantly more frequent collocate of “M5” in the blog posts, refers to aspirations or intentions and is realised by various word forms in relation to air traffic. Some of these uses have a political or military dimension (e.g. *policy*, *target*, *missions*). The tag also includes personal intentions, as illustrated in Examples (7) and (8). The examples use the “X7+” collocates in different ways to express disagreement with airport security checks (see also Section 4.4.2.1). Both reflect negatively on the surveillant landscape at the airport and suggest discomfort felt during the check (see “hissy fit”, “not reasonable”). These individuals clearly object to the “reading” capabilities of the surveillant landscape, suggesting that the security checks are invasive. The discomfort seems to relate to feelings of exposure and powerlessness, i.e. the social relationships, or interaction orders, set up in these situations (see Section 5.4). In (7), the disagreement with the checks is expressed with irony about an exaggerated mind-reading scenario that would facilitate more efficient checks (see the rhetorical questions). Example (8) is a newspaper quote of a pilot who refuses to undergo security checks. The blogger uses this quote to reflect on the strip search procedure.

- (7) I had a hissy fit over full-body scanners being installed in airports as a knee-jerk reaction to the underwear bomber - so that led me to thinking - why bother with the tardis -like machines that x-ray the heck out of you? Why not have machines that quickly read a passenger’s mind? Shove us in booths with flashing lights and zapping sounds (for special effects) and read our minds to see if we have evil *intentions* to blow up **planes**.
(wordpress_outfile231; thinkingshift, 2010)

- (8) “It’s not reasonable when you walk into the airport, and just because you *want* to fly on an **airplane** that they should strip search you, or physically put their hands on your crotch or feel your body from top to bottom,” he said.

(blogspot_outfile093; CFC, 2010)

Another tag relates more explicitly to the interaction order in surveillant landscapes: the “S7.1+” (“Power, organising”) tag is found to be significantly more frequent in the blog posts. Collocates with this tag mainly refer to institutional power (roles), e.g.: “Transportation Security Administration”, “director”, “coordinator”, “headquarters”. While these mostly relate to official institutions, they can also refer to terrorist organizations (e.g. “the *leader* of the biggest **plane** hijack in history”).

This section has argued that the “M7” and “M5” tags are useful entry points into the surveillant landscapes depicted in the Surveillance Blog Corpus. The collocates point to a wide range of contexts, including surveillance against civilians, on college campuses, in relation to financial crime, and, for “M5”, in the context of airports. The examples raised some concerns in relation to surveillant landscapes, and feelings of powerlessness, which are particularly relevant to interaction orders (also see Section 5.4). At the same time, the analysis has reflected the heterogeneity of the corpus with references to flight missions and promotional descriptions of surveillance.

5.3.2.2 Co-occurrences of semantically tagged M nodes

The aim of this section is to take one more step towards gaining a more general overview of surveillant landscapes in the corpora by looking at the “M” tag as a whole. First I provide the results for the collocates at the level of subcategories before moving on to the most general analysis of both nodes and collocates at the representing the major discourse domains.

Table 5-6: Semantic collocates of “M” nodes with the highest effect sizes (excluding “unique” collocates)

Rank	Blog posts	Journal articles
1	X6 (“Deciding”)	Q1.2- (“Paper documents and writing [-]”)
2	O4.6 (“Temperature”)	H4- (“Residence [-]”)
3	O1.2- (“Substances and materials generally: Liquid [-]”)	N3.6 (“Measurement: Area”)
4	O1.3 (“Substances and materials generally: Gas”)	K6 (“Children’s games and toys”)
5	G3- (“Warfare, defence and the army: weapons [-]”)	S7.4 (“Permission”)
6	A8+ (“Seem/appear”)	A4.2- (“Particular/general; detail [-]”)
7	O1.2 (“Substances and materials generally: Liquid”)	A1.8- (“Exclusion”)
8	O4.6+ (“Temperature [+]”)	A11.2+ (“Importance: Noticeability”)
9	N3.5+ (“Measurement: Weight [+]”)	E1 (“Emotions: General”)
10	N3.4- (“Measurement: Volume [-]”)	X5.1+ (“Attention [+]”)

The semantic tag collocates with the top ten biggest difference for each corpus are listed in Table 5-6, as the plot of all significant results is too large to print here. The table provides an initial impression of the salient patterns for each corpus. The top semantic collocates of the “M” nodes most peculiar to the blog posts can be broadly summarised as describing objects and events in terms of their size, material, temperature, etc. or to refer to descriptions themselves as in the “Deciding” tag (“X6”). These findings to some extent echo those for the “M7” node (Section 5.3.2.1). The decision collocates are mostly realised by word forms of ESTIMATE and JUDGE. In Example (9), the tag is applied to describe the approximations of the surveillant technique of crowd-counting. The example questions the effectiveness of this surveillance technique, arguing that it is “inexact” and “far more art than science”.

- (9) “[...] Still, it is difficult to use a photograph taken at a particular point in time to count people who are **moving** and **shifting**, and *estimates* are always going to be quite inexact -- far more art than science.”

(wordpress_outfile248; Zeese, 2013)

In the journal column, nearly all of the instances of the top rank (the negated “Paper and documents” tag) are realised by the word form *undocumented*, e.g. in “capture *undocumented*

migrants during their **crossing** of the border” (2008-05-2-02; Ceyhan, 2008). This pattern relates to the theme of “Identity-based Surveillance” of vulnerable groups (see Section 4.4.1), which suggests an interaction order in which the undocumented migrants are powerless.

To gain a wider overview of the semantic differences, we can group not only the “M” tag as a main category, but also the co-occurrences. When all tags are considered at the top level, only the 21 “major discourse fields” (Archer et al., 2002, p. 2) are part of the comparison. Accordingly, the results can easily be displayed in a small plot (see Figure 5-13).

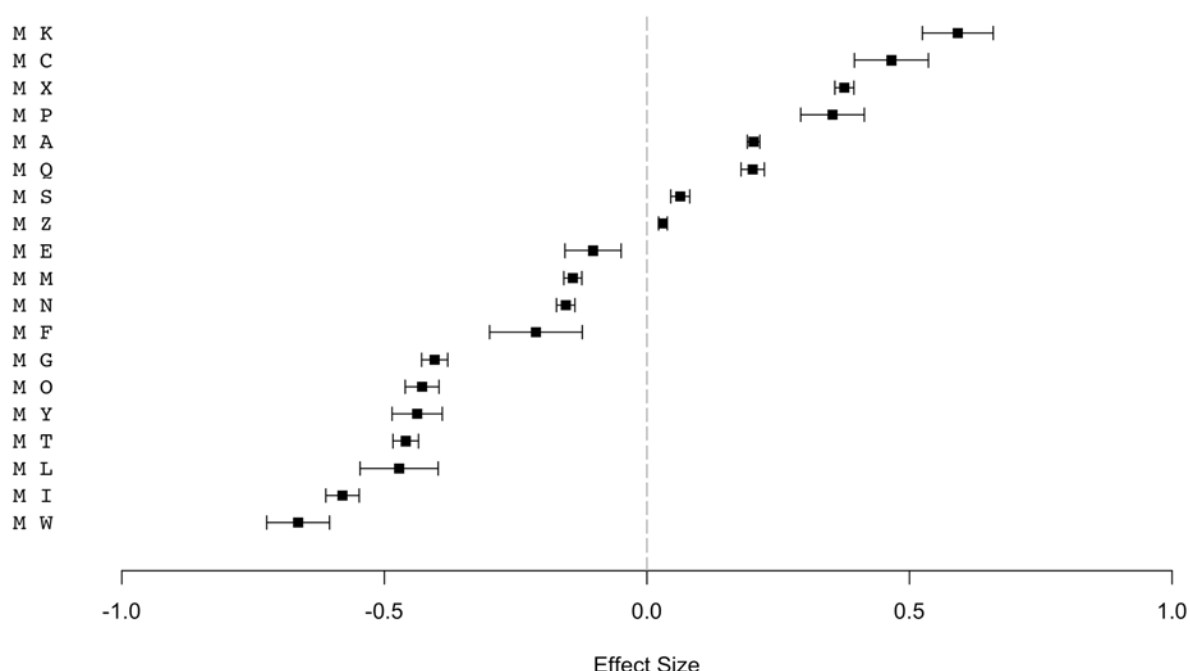


Figure 5-13: Co-occurrence comparison of main semantic tags (blog posts left; journal articles right)

With the conflation of the subcategories the results not only become more compressed, but also more abstract. In the blog posts, co-occurrence of “M” with the tags “The world & our environment” (“W”), “Money & commerce” (“T”), “Life & living things” (“L”) and “Time” (“T”) are most different from the journal articles. The “W” collocates include tokens like *world*, *earth*, *satellite* and *mountains*. The “T” tag (“Money and Commerce”) has featured in previous

plots, particularly in relation to money and debts, although it also covers employment. Whereas the journal contains a special issue dedicated to workplace surveillance, 1(2), the collocates of this category in the blog posts (such as *contract*, *budget*, *corporation*, *tax*) may be frequent there due to their applied relevance in relation to business affairs and financial news. Some blog posts are actually authored by corporations. Example (10) originates from such a commercial post, which follows the format of a press release in announcing the company's success in "securing a contract" for "CCTV surveillance systems".

(10) Premier security integrator 2020 Vision Systems has won the trust of Northumberland, Tyne and Wear NHS Foundation Trust; securing a *contract* to maintain their **estate** wide CCTV surveillance systems.

(blogspot_outfile003; McIntyre, 2012)

Concordance 5-8 shows that collocates from tagged as "L" refer not only to life and being alive (lines 1, 4 and 9, *life*; line 8, *live*; line 10, *alive*), but also its semantic opposite, death (line 2, *passed away*; line 3, *killing*, line 7, *deadly*). Possibly the blog posts are more concerned with the lives of particular people (e.g. line 2) and the discussion of news and current events (line 3), (see research on local coverage in blogs, Section 2.5.2; Gordon & de Souza e Silva, 2011). In addition, this sample of concordance lines contains words tagged as "L2" ("Living creatures generally"), which are metaphorically used for describing people, such as *flocks of people* (line 5) and *a wing nut like Aquino* (line 6). These rather colloquial expressions would be less likely to appear in the research articles of the S&S Corpus.

[1]	to start a new <i>*life*</i>	elsewhere	. In spite of the surveillance
[2]	to name a few. Earlier	this	year, Janet Veal <i>*passed*</i> <i>*away*</i>
[3]	Oklahoma City <i>*killing*</i> 168, one	left	legwas found in the rubble.
[4]	and community <i>*life*</i> . My community	approaches	an archive that has documentation
[5]	<i>*flocks*</i> of people were to	leave	for another social network, there
[6]	a <i>*wing*</i> nut like Aquino	access	to the secretive corridors of
[7]	<i>*deadly*</i> this can be, if	cyclotron	resonance is intentionally used to
[8]	<i>*live*</i> long enough to see	this	horror completely unfold, however,
[9]	NSA and DHS. <i>**Towards</i> the	end	of his <i>*life*</i> , even Dr.
[10]	<i>*alive*</i> who were born before	this	cataloging of our medical

Concordance 5-8: Sample of the “M” tag co-occurring with the “L” tag in blog posts (10 out of 3,312 examples)

According to Figure 5-13, time references occur more frequently with the “M” tag in blog posts compared to journal articles. The blog register clearly allows a much more immediate response to events, making time an important aspect of blogs (see Myers, 2010, Chapter 5). This immediacy is reflected in Example (11). Underlined tokens represent “T” tags in this passage, most of which are deictic markers. The time frame is important for this post: the example explicitly states that the post is written in response to “breaking news” published since the last post on the topic only “two weeks ago”. Bold tokens are realisations of the “M” tag and its subcategories. Two mobility and place references refer to the physical world (*airport*; *passenger*). They relate to the example of the pilot refusing airport security in Example (8), Section 5.3.2.2, which originates from the same post. The remaining place tags are not physical references. *Backdrop* is a metaphorical use. Then there are various “virtual” place references: the instances of *here* (tagged as “M6”; “Location and direction”) at the end of the example are hyperlinked in the original and take the reader to the blogger’s “past posts”.⁹ So, these links encourage the reading of the bloggers’ other posts. They therefore make other areas of this blog’s a virtual landscape more visible, like surveillant landscapes encourage the reading and writing by visitors. These virtual versions of spatial deictic markers illustrate how “bloggers use the language of space to construct” the blogosphere (see Section 2.6.1; Myers, 2010, p. 50).

⁹ The original format of this example also contains bold and multiple font colours as such formatting is not reproduced for any of the examples.

(11) I just posted on **this** subject two weeks ago...but more breaking news has come to my attention. To give readers a **backdrop**, I'm first going to use some of the intro I wrote for that last post [...] Most important to understand is that these machines essentially allow **airport** security to see through your clothing, producing images of digitally naked **passengers**. Now, I don't want to rehash all that I have written on this subject before because there's a lot I want to get to today, so to find out most everything you need to know about these machines and their privacy implications (among other issues with them), check out my article "The Politics of Fear and Whole Body Imaging" (from January 2010), or check out some of my past posts on the subject, **here** , **here** , **here** , and **here** .

(blogspot_outfile093; CFC, 2010)

Some of the results from this abstract comparison of the major discourse fields seem counterintuitive at first sight, specifically on the side of the journal articles. For example, the most distinct semantic co-occurrences of the "M" tag on the side of the journal articles are the "K" ("Entertainment, sports & games") and the "C" ("Arts & crafts") tags. However, this makes sense when considering that the journal contains special issues on related topics: 7(2) *Surveillance, Performance and New Media Art*, 11(4) *Surveillance and Sport* and 12(3) *Surveillance, Gaming and Play* (see Section 3.2.1.1). At the same time, the nature of the *USAS* tagging procedure means that not all references to "C", for example, strictly refer to arts and crafts. Most importantly, concordance lines show that instances of *camera* are counted as part of the arts and crafts category. In Chapter 4, the KKWs *camera* and *cameras* of the S&S Corpus were also placed in this *USAS* category (see Section 4.3.2). Because CCTV cameras are widespread surveillance gadgets, the high frequency of this category is not surprising in the journal. It is noteworthy that these references are not equally frequent in the blog posts. This finding suggests that CCTV cameras are of a higher importance for the surveillant landscape of the journal. It may also indicate that the Surveillance Blog Corpus is more heterogeneous and less specialised in its discussion of surveillance.

So, this section has moved to the most abstract level of the comparison. The section started by comparing co-occurrences of the "M" tag as a whole with semantic tag subcategories

across the two corpora. That comparison reflected some of the findings for the “M7” tag in Section 5.3.2.1. The second part of this section showed that in the most abstract terms, the surveillant landscape in the blog posts is more concerned with general references to the world and environment, time and life and death. I have suggested that these findings are related to the blog posts responding more quickly to current events and financial news or business affairs. As many posts quote from news sources, it is likely that the news values of negativity and superlativeness (see Section 2.5.3) play a role in the blog posts, for example when they refer to conflicts.

5.4 Interaction orders of surveillant landscapes in blog posts

This chapter argues that the social dimension of surveillant landscapes (the interaction orders) is a crucial element of the theoretical framework. I propose that interaction orders are the most appropriate aspect to focus on when following a corpus linguistic approach. The medium of the blog post is particularly suitable for an interaction-focused analysis (see Section 5.1), because the act of posting about the surveillant landscape is an interaction with the landscape in itself. Among the semantic tags, two categories are the most relevant for the analysis of interaction orders: in particular the “S” tag (“Social actions, states and processes”), examined in Subsection 5.4.1 and the “E” tag (“Emotional actions, states and processes”), discussed briefly in Subsection 5.4.2.

5.4.1 Social aspects

The general comparisons of previous sections have highlighted the role of some social tags, such as the “Competition” “S7.3(+)” tag (Figure 5-11, Section 5.3.2.1) and the “S7.1+” power tag (Figure 5-12, Section 5.3.2.1). This section restricts the analysis to co-occurrence pairs of

“M”-tagged nodes and “S”-tagged collocates. Table 5-7 shows the first level of the “S” subcategories with their labels and example nodes.

Table 5-7: The first 20 distinct realisations of each “S” subcategory in the Surveillance Blogs Corpus (ordered alphabetically)

Tag	20 examples from the blog posts
S1 (“Social actions, states & processes”)	accessible, anthropologist, anthropologists, behaviour, ceremonies, formal, habits, initiation, interference, interfering, reasonable, respond, roundtable, sharing, social, strong, tough, traditional, treating, visit
S2 (“People”)	and, ¹⁰ character, characters, child, children, girl, human, identities, identity, individual, individuals, kids, man, nonmembers, other, people, peoples, person, persons, woman
S3 (“Relationship”)	acquaintances, associates, encounter, friend, friends, homosexuality, intimacy, love, meet, other, partner, partners, relationship, relationships, romantic, separated, sexually, that, the, way
S4 (“Kin”)	ancestors, clan, daughter, divorced, families, family, father, foster, grandchildren, grandmothers, have, husband, kids, married, mother, mothers, parents, son, twin, twins
S5 (“Groups & affiliation”)	alone, branch, communities, community, fbi, federal, following, fraternal, group, groups, institutions, member, members, organization, personal, society, sub-groups, subgroups, together, tribe
S6 (“Obligation & necessity”)	duty, essentially, free, have, must, necessarily, necessary, need, needed, needs, obligation, obligations, ought, patriot, precondition, responsibility, should, supposed, to, waived
S7 (“Power relationship”)	allowing, authorize, competing, consent, control, custodian, governed, heading, managed, manager, ordering, organizing, permission, promotion, respect, respectful, right, rights, senior, supervision
S8 (“Helping/hindering”)	assist, champion, encouraging, help, in, of, patron, patrons, preservation, prevent, prevented, preventing, prevention, promote, protect, protecting, protection, protections, spite, undermined
S9 (“Religion & the supernatural”)	advent, catholics, demonic, goyim, jews, magic, muslims, new, old, protestants, religion, religious, ritual, rituals, roman, sacred, satanic, soul, spiritual, testament

¹⁰ *And* is tagged as “S2” when it forms part of a multiword unit. Specifically, it seems to be part of the pattern *X and other Y*, which appears hundreds of times in a concordance of all (over 17,000) S2 nodes. Some of these instances refer to people such as *Felix Rodriguez and other anti-Castro Cubans*, but many seem unrelated (e.g. *Africa and other parts of the world*). This chapter does not focus on the S2 collocates, because even the legitimate people references tagged as “S2” form a rather too vague or inclusive collection. This tag does not seem useful for the present purpose of exploring the interaction order in the given corpus.

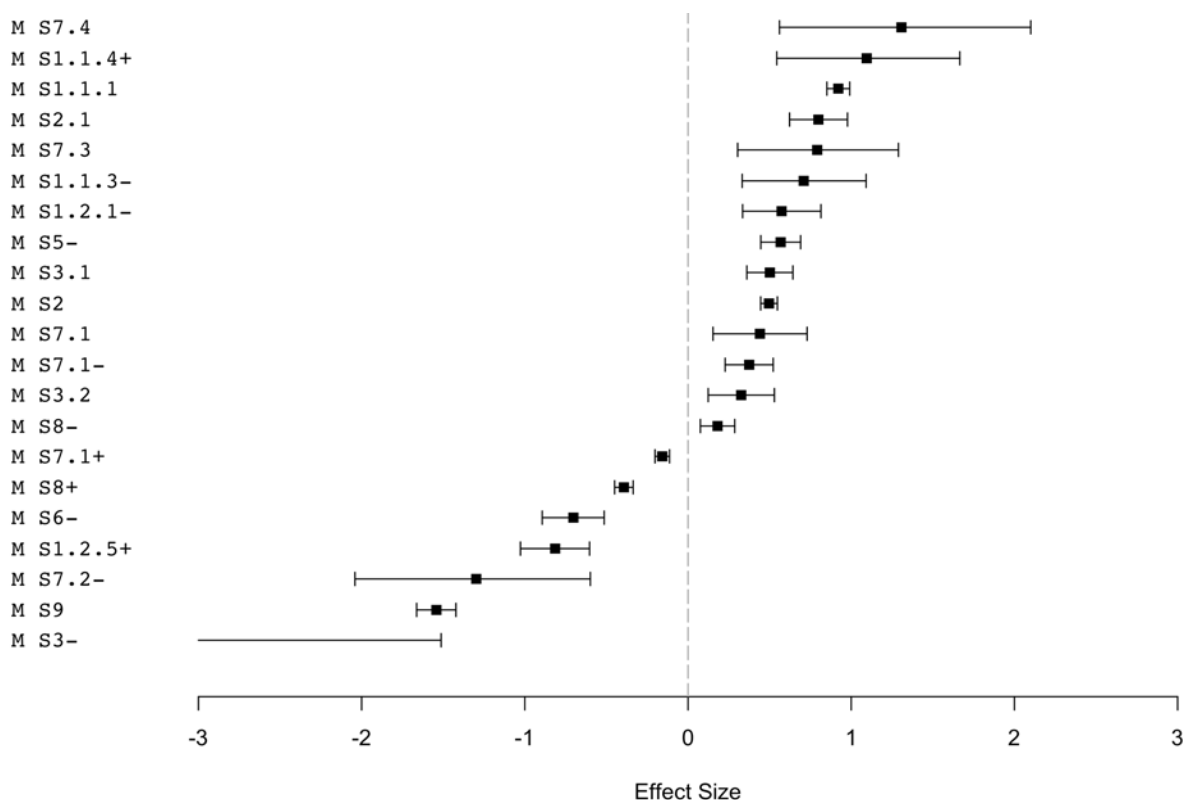


Figure 5-14: Co-occurrence comparison of the “M” semantic tag and “S” (“Social actions, states and processes”) subcategories

Figure 5-14 shows that “S9” (“Religion & the supernatural”) is the most frequent collocate of the “S” subcategories in the blog posts compared to the journal articles.¹¹ Several examples of metaphorical language use are categorised as religious, including “embodies a new *spirit* of **national** development”. In other cases, the words refer to religious concepts, like *Satan* and *angels*, as shown in (12). This blog post probably fits into Biber and Egbert’s (2016, p. 116) category of RELIGIOUS BLOGS, a sub-register of opinion blogs that are characterised by a “detached” linguistic style.

(12) *Satan* and his hideous **fallen** *angels* want to torture us for all eternity.
(blogspot_outfile252; Dwight, 2009)

¹¹ The unique tag “S3-” (“No personal relationship”) only occurs seven times, all realised by the collocate *split (up)*.

The next most different “S” subcategories that co-occur significantly more frequently with the “M” nodes in the blog posts are “(Dis)respect” (“S7.2-”), “Toughness” (“S1.2.5+”), “No obligation or necessity” (“S6-”), “Helping” (“S8+”) and “Power, organizing” (“S7.1+”). The last three of these categories appear particularly relevant to the interaction orders of surveillant landscapes. Their comparatively low effect sizes and narrow confidence intervals indicate that these collocates are highly frequent in both corpora, but slightly more frequent in the blog posts. A very large difference between the two surveillance corpora for the power tag would be rather surprising, given that power and control are fundamentally related to surveillance (see Section 5.4.1.3; Lyon, 2007, p. 15).

Subtle differences do seem to exist in the polarity of the semantic tags. For both “S8” and “S7.1”, the positive subcategories appear on the blog side of the results: “S8+” (“Helping”) and “S7.1+” (“In power”). By contrast, the general or negative subcategories of these tags are found on the side of the journal: “S8-” (“Hindering”), “S7.1-” (“No power”) and “S7.1” (“Power, organizing”). The themes of obligation and power relate to the questions of responsibility and rights (as can be seen from the examples in Table 5-7), to which the definition of interaction orders refers (Jones, 2017; quoted in 5.2.1). The notions of “helping” and “hindering” (tag “S8”) can imply evaluation of the surveillant features and therefore also point to an interaction between the watched and the watcher. Based on the relevance to surveillant landscapes, the analysis places particular focus on the co-occurrences of the “M” tags with “No obligation or necessity” (“S6-”, Section 5.4.1.1), power relationships (“S7”; Section 5.4.1.2) and the concepts of “helping” and “hindering” (“S8”; Section 5.4.1.3).

5.4.1.1 No obligation or necessity (“S6-”)

In the blog posts, mobility-related lexis (the “M” tag) co-occurs significantly more often with items at the negative semantic pole of the “obligation and necessity” category. So, patterns tagged as expressing little or no obligation are more often found with the “M” nodes in the Surveillance Blog Corpus than in the S&S Corpus. Example (13) illustrates this “dispensable” meaning of the “S6-” tag (realised by *waived*).

- (13) We might say that, while my conversation is private even if I am speaking loudly on a public **street**, I have *waived* my privacy rights (see Thompson, 302). In that case, the critical issue is to say under what conditions I have waived my rights.
(blogspot_outfile002; Lenhart, 2013a)

The example has an academic dimension to it (see the in-text citation), because it originates from an educational blog, set up as part of an “information ethics” class, and contributes to a discussion about the ethics of privacy. The (“M”, “S6-”) pattern is here used as part of a thought experiment: the inhabitant of a surveillant landscape “waives” their privacy rights by speaking in a public space. Similarly, the “S6-” tag is realised by the word *exemption* in (14) from a legal blog that often discusses surveillance and data protection-related legislation. This example problematises a data protection exemption for national security purposes. Here, the “M” tag is realised by *national*, which refers to a more general concept than a concrete public space like *street* in (13).

- (14) One can also see that the fact that this Code of Practice is offered as a protection implies that the Government is anticipating the continuation of an unchanged Section 28 **national** security *exemption* in the Data Protection Act.
(typepad_outfile017; Amberhawk, 2016)

One of the most frequent realisations of the “S6-” collocate, however, is *free*. This collocate is used in a variety of contexts, including commercial in (15), freedom from prison in (16) and freedom of speech in (17). Examples (16) and (17) have a decidedly political dimension. Both clearly contribute to a critical public discourse on surveillance and problematise an interaction order in which the powers of the state go too far. Accordingly, these examples would fit well into the next subsection on power. However, the obligation and necessity tags have their own functions here. The blogger in Example (16) complains about the lack of liability that “REAL CRIMINALS” face while “DECENT CITIZENS ARE TREATED AS CRIMINAL SUSPECTS” (emphasised with capitals). Example (17) expresses a similar point but stresses the need for citizens’ freedom of speech (and press).

- (15) **Apex** CCTV offers *free*, lifetime technical support on our products
(blogspot_outfile174; Pad, 2010)
- (16) WE LIVE UNDER A SOCIALIST REGIME WHERE DECENT CITIZENS ARE
TREATED AS CRIMINAL SUSPECTS, YET REAL CRIMINALS **WALK** *FREE*
BECAUSE THE AUTHORITIES HAVE BEEN SO ENFEEBLED BY
BUREAUCRACY AND LEFTWING DOGMA.
(blogspot_outfile215; “I AM AN ENGLISHMAN,” 2009)
- (17) [...] in the spirit of universal transparency all citizens should be *free* to publish and
access information about corruption and other crimes without fear of prosecution.
(wordpress_outfile246; Undercover1, 2015)

Another frequent “S6-” collocate of the “M” tag is *NSA*. This appears to be a case where the tagger’s matching of a type with its lexicon reaches a wrong conclusion. The concordance lines suggest that the abbreviation is used in the Surveillance Blog Corpus to index the US National Security Agency as in (18). This usage does not fit into the category of having no “obligation or necessity” although it has been tagged as such. Nevertheless, the concerns voiced in this example are similar to those in previous examples about the magnitude of power of a government institution.

(18) Moreover, the *NSA*'s illegal **access** to our minds via the EEG heterodyning of the neural pathways of our own brains, must also be considered a type of electronic trespass, since for all intents and purposes the brain can be likened to an organic computer [...]
(blogspot_outfile008; Marino, 2014)

USAS appears to tag the abbreviation *NSA* as a short form of the phrase *No Strings Attached*, which fits the negative end of the semantic field of obligation and itself co-occurs with the “M” tag, as shown in (19). The blog from which (19) originates, campaigns for privacy in the healthcare system of British Columbia (BC), Canada.

(19) 100+ reasons the BC Liberals must **go**, *No* *Strings* *Attached*: Laila Yuile on politics and life in B.C.
(blogspot_outfile194; Searcher, 2016)

The issues with the semantic tagging of context-specific terms (like *NSA*) notwithstanding, the examples in this subsection suggest that a variety of perspectives are at work in the interaction orders of surveillant landscapes. Customers of surveillant equipment enter commercial interaction orders with their supplier. Citizens enter different interaction orders with others around them in the surveillant landscape (e.g. when speaking on a public street) and with governmental institutions when publishing contentious opinions. Patients provide their personal information to healthcare professionals and may not know who it is shared with. It takes effort to question the surveillant practices and thereby question the existing interaction orders, as some of the quoted bloggers do.

5.4.1.2 Helping (“S8+”) vs hindering (“S8-”)

The interaction orders of surveillant landscapes, so the social relationships set up in the environment, are shaped by the perception that surveillance is a double-edged sword. A balance has to be found between its benefits (such as security) and costs (both financial and in terms of invasions of people’s privacy or loss of freedoms). The “S8+” tag, which is significantly more frequent in the blog posts, may contribute to this discussion as I illustrate with the analysis of three categories: (i) “promotional” language in favour of surveillance, (ii) concerns about surveillance and (iii) perspectives on helping. These categories have emerged from a subset of the concordances for the co-occurrence pair (“M”, “S8+”). I decided to focus on the 1,251 concordance lines in which the “M7” (“Place”) tag is used (out of 5,477 total lines), because (i) the relevance of the “M7” tag has been demonstrated in Section 5.3.2.1, and (ii) the pair (“M7”, “S8+”) is itself significantly more frequent in the blog posts.¹² My focus was on salient examples and patterns rather than on quantitative results because examples can match multiple categories. So the categories are not mutually exclusive and no quantitative claims are made about the spread of these elements in co-occurrences of the “M7”/“M” tags. The analysis in this section reiterates that meaning is contextually-bound.

(i) “Promotional” language in favour of surveillance

The first category includes examples where the “S8+” tag describes the benefits of a surveillant technique. Example (20) contains two co-occurrence pairs of the “M7” node with “S8+”. The first instance specifically describes the benefits of surveillance cameras in the area. The second one generalises this to “advanced techniques” that facilitate the surveillance of remote locations

¹² The “S8+” tag is not shown in the main “M7” comparison (Figure 5-11, Section 5.3.2.1), because the effect size for this pair is below the threshold of one that has been applied to that plot.

without security guards or policemen onsite. While only the two collocates marked with asterisks are tagged as “S8+”, the context further enhances the positive evaluation through the intensifier *extremely* and the adjective *efficient*, which suggests a commercial dimension. So, the extract frames passers-by as unwanted (*trespassing*) and presents the surveillant practices as useful.

(20) These cameras are extremely **helpful** in monitoring high risk **areas**, and remote places that are vulnerable to trespassing. Advanced techniques eliminate the need for human presence at these **sites** and **serve** to make the surveillance process more efficient.

(blogspot_outfile102; JQ, 2005)

(21) New York police are considering plans to place hundreds of video cameras throughout the **city** to **help** fight crime and combat terrorism, the New York Police Department said on Monday.

(blogspot_outfile287; BIGMART, 2005)

In a similarly positive way, (21) ascribes the quality of “helping to fight crime and combat terrorism” to the planned camera scheme. However, this is a quote by the police department who are introducing the scheme rather than an overt evaluation by the blogger. Moreover, the wider context shows that this example is part of a news article pasted into a newsletter on the blog *BIGMART’s Privacy solution site* under the heading “Bad News”. The context is therefore important.

The concordance lines of “M7” and “S8+” contain several examples of “promotional language”. These tend to appear in quotes that promote the values and/or practices of institutions. Example (22) refers to the mission statement of the FBI, which is quoted to evaluate surveillant actions positively, in national defence “against terrorist and foreign intelligence threats”. Yet, this and other long quotes from security services and similar institutions on this blog are framed critically; the title for this section is “BIG BROTHER’S CURRENT GLOBAL INFRASTRUCTURE”. Despite the positive quote, the context again contains criticism, with

this section title clearly suggesting an Orwellian interaction order of the surveillant landscape put in place by the state.

- (22) Commitment to these values and standards ensures that the FBI effectively carries out its mission: Protect and defend the United States against terrorist and **foreign** intelligence threats; *uphold* and enforce the criminal laws of the United States [...]
(blogspot_outfile030; Guillaume-sam, 2007)

A similar point is made in (23), which also relates to the FBI, but discusses a specific surveillant practice: the control of library book borrowings. Here the promotional language is not pasted as a big chunk, but indirectly quoted (“protecting the country from foreign spies”) and marked by the phrase *their concern for*. Criticism is indicated by the scare-quotes around the verb *justified* and explicitly explained in the following sentence (“What’s problematic about this”).

- (23) And they “justified” this claim on the basis of their concern for *protecting* the **country** from foreign spies. What’s problematic about this isn’t merely that national security couldn’t be a legitimate ground for such behavior, it’s also that the behavior doesn’t actually protect national security.
(blogspot_outfile002; Lenhart, 2013b)

Other promotional language does not clearly justify surveillant practice or necessarily relate to them. The “S8+” tags are frequently realised by phrases referring to organisations (e.g. *Secret Service*, *International Institute for Democracy and Electoral Service*, *Google Services*, *technical support*) and institutional roles (e.g. *Homeland Security’s first Assistant Secretary*, *LDP Administrative Assistant*). These instances are similar to the institutional power roles in Section 5.3.2.1 tagged as “S7.1+” (e.g. *director*), with the difference that the institutional titles tagged as “S8+” here contain “supporting”, “service” or “assisting” elements.

(ii) Concerns about surveillance

The following examples clearly voice concerns about surveillance. In these cases, the “S8+” tag does not positively evaluate the surveillance practices but refers to the context. Example (24) illustrates a comparative discourse of privacy matters from the blog campaigning for patients’ privacy rights (quoted also in Example 19). By handing out flyers in public, the blogger engages in a dialogue with other passers-by, thereby learning about other perspectives on the global surveillant landscapes of hospitals. The contrast with “the really good rules for protecting patients privacy” in “European countries” implies that local laws are insufficient (which is made explicit in the wider context).

(24) The European’s [sic] discuss how the different **countries** *protect* people’s privacy. Some European countries appear to have really good rules for protecting patients privacy.

(blogspot_outfile194; Searcher, 2016)

A more specifically location-related illustration of a privacy issue is discussed in (25) from the weekly news digest of the blog *Privacy News Highlights*. In this example, the “S8+” tag is negated (“isn’t *protected*”), highlighting that the patterns in which tags appear are important for interpretation.

(25) “Right now we protect health data, we protect financial data, we protect kids’ data, but **location** isn’t *protected*,” said [...]

(wordpress_outfile171; privacynewshighlights, 2014)

Not all surveillance concerns are restricted to privacy. Example (26) originates from what appears to be an activist blog campaigning against the London 2012 Olympic Games and particularly the state control and policing measures introduced for this mega event. Olympic

surveillance is picked up by surveillance studies scholars, e.g. in the S&S Special Issue 11(4) *Surveillance and Sport*. The blogger argues that the Olympic games serve as an excuse for the state to intensify surveillance measures. The realization of the (“M7”, “S8+”) co-occurrence occurs at the end of this example where the verb *supporting* does not refer to the surveillant practices but to people in need (including migrants) with sleeping places. In calling for support, this example suggests that the sports mega event is accompanied by forms of identity-based surveillance (see Section 4.4.1).

- (26) The ‘clean up’ of cities in London and Calais in the name of keeping the Olympic industry and its spectators secure, is in reality an opportunity for reinforcing the deployment of draconian surveillance systems and arsenal of security measures to protect the privileged and their profit-making.

[...]

Call out for squatters: *supporting* people find **places** to sleep would be much appreciated at the moment!

(blogspot_outfile241; Sevrino, 2012)

A final example illustrates another specific issue related to surveillance concerns; activism against ID cards. While (27) refers to legislation in the US, Chapter 6 explores the theme of ID cards in the context of British newspaper coverage. As with previous examples grouped in this category, the positive “S8+” tag is here accompanied by negative lexis (“trick Americans into”). This example specifically warns of the dangers of a surveillant landscape, using the terminology *total surveillance society*. Chapter 6 shows that the bigram *surveillance society* is also frequently used in the British context, appearing in *The Times* frequently from the 2000s onwards.

- (27) Congress is trying a new flimflam to trick Americans into *supporting* an **International** ID and a total surveillance society. Now they are pretending to REPEAL the REAL ID Act and replace it with the PASS Act (S.1261).

(blogspot_outfile316; Stephens, 2015)

(iii) Perspectives on “helping”

The concordance lines suggest that the function of (+/-) polarity of the “S8” tag depends on the perspective provided in the context. In (28), the “S8+” tag co-occurs with *terrorist organization*. This extract reports the reason for a court sentence, showing that the object of the “S8+” tag does not have to be an entity that the writer supports.

(28) In July, the Berlin Superior Court sentenced Fatih I. (28) to three years and six months imprisonment on two counts of *supporting* a **foreign** terrorist organization as well as fraud.

(wordpress_outfile111; News Agency News247WorldPress, 2016)

The “S8+” collocate in (29), *defense*, although tagged positively, implies criticism of the surveillant activity in a similar way to the scare-quoted *justified* in Example (29). In elaborating on the reported dialogue, the blogger uses scare-quotes and comments on this usage meta-linguistically (“It has to be in quotes”). This positions surveillance as a discourse topic worthy of debate, even though this experience is markedly more negative than that of the blogger discussing patient privacy with passers-by in (24). At the same time, the mention of the “recent revelations of the NSA” in Example (29) provides evidence that the Snowden’s “revelations” received attention across various discourse domains beyond academia (cf. the discussion of Volume 13 in the S&S Corpus, Section 4.4.2.3).

(29) The topic, of course, was about the recent revelations of the NSA snooping on American’s cellphone calls, international calls, world leader’s calls, etc. And the comment was a defense of this government snooping activity: the “Nothing To **Hide**” *defense*. It was a very frustrating and disappointing “conversation.” It has to be in quotes because it wasn’t a conversation, not a discussion, not even a debate.

(wordpress_outfile212; Coronare Modestus Faust, 2013)

The list of examples for the “S8” subcategory as a whole in Table 5-7 contains various forms of the word *prevent*. This word is not featured in the previous examples of this section because

it is tagged as “hindering” (“S8-”) rather than “helping” (“S8+”). However, with a topic like surveillant techniques these two concepts seem to be closely related. In Example (30), the two tags co-occur (*helping* is tagged with “S8+” and *prevent* with “S8-”), because the prevention – in this case of misidentification, in other cases perhaps of crime – is seen as positive.

(30) It scans more of the fingertip’s surface **area**, *helping* to **prevent** misidentification.
(blogspot_outfile221; tan, 2004)]

As indicated at the beginning of the section, the “hindering” tag co-occurs significantly more frequently with “M” tags in the journal articles as compared to the blog posts. Like the difference in “S8+” tags, the difference is not vast (as seen in the low effect sizes). Both corpora contain many co-occurrences of the “M” with the two “S8” subcategories, supporting the idea that notions of “helping” and “hindering” are salient in surveillance discourse. Realisations of the “S8-” tag in the journal’s co-occurrence patterns of with the “M” tags are particularly shaped by the notion of resistance, which is the topic of Special Issue 6(3) (see Section 4.4.2.2). As a theoretical term, *resistance* can be found in general comments, for example co-occurring with metaphorical place references like the “territory of resistance” in Example (31).

(31) This is the **territory** of *resistance*, civil disobedience or even criminal subversion of authorities’ surveillance.
(2014-12-1-12; Stoddart, 2014)

These examples only scratch the surface of the patterns in the journal articles, which mainly serve as a reference corpus in this chapter. The apparent reason for concepts like resistance appearing more frequently with the “M” tags in the journal articles than in the blogs is that the

surveillance studies scholars tend to analyse and comment on activities of surveillance dissent, while the blog posts may form part of these activities¹³.

Overall, this subsection has illustrated that the semantic “helping” collocates of “M” tags fulfil various functions. Although the collocates themselves are positive, the co-occurrences may not always correspond to agreement with surveillant practices. In those instances that clearly refer to surveillance, the context may contain evidence for negative evaluation. In particular, elements of “promotional language” may be used to strengthen commentators’ criticism of institutional surveillance. Therefore, the section has proposed that the blog posts depict a complex picture of who does and does not benefit from particular surveillant landscapes. The analysis also suggests that the challenges of analysing these perspectives on surveillant landscapes essentially boil down to the general problem that evaluative language has many forms (see Hunston, 2011). So, the polarity depends on the perspective of the speaker/writer.

5.4.1.3 Power relationships (“S7”)

Control and power are central aspects of surveillance. Depending on the context, forms of power and control can be found at a small or large scale and with benign or bad intentions. As Lyon (2007, p. 15; emphasis mine) points out, “it should not be imagined that the influence, management or *control* is necessarily malign or unsocial, despite the frequently negative connotations of the word ‘surveillance’”.

One group of power collocates in the blog posts refers to leadership positions and institutions such as *leaders*, *chief*, and *management*, similar to the “S7.1+” examples discussed

¹³ Some work in surveillance studies may be considered “activist”, in particular the artistic presentations (which were excluded from the S&S Corpus, see Section 3.2.1).

in 5.3.2.1. Example (32) illustrates the use of institutional collocates. Although (32) is clearly critical of the use of the National Military Command Center in relation to its handling of the 9/11 attacks, this criticism is conveyed not by the “power” collocates themselves (*command*, *coordinated*) but their context (“the pathetic official version”).

- (32) The second crash was shown on live television at 9.03 and this is the Defense Secretary sitting two hundred feet from the National Military *Command* **Center** from where hijack responses are *coordinated*. If you only believe the pathetic official version that Pentagon officials were given twelve minutes’ warning of the plane heading their way (unlike Cheney), why would no one tell the Defense Secretary before he felt his ‘jarring thing’?

(blogspot_outfile192; Hissil, 2009)

A much more positive and empowering use of a power-tagged co-occurrence is given in Example (33), where civilians hold the control. It suggests a different interaction order from the surveillant landscapes characterised by state control as represented in Section 5.4.1.2 on “helping” and “hindering”. The wider context of the post explains that residents are introducing surveillant measures in order to reduce actions such as drug dealing and prostitution in their neighbourhood. Surveillance is framed as a way to “take back control”, so these residents actively improve their own role in the interaction order.

- (33) Gray says Hilltop residents are supplementing cameras with other measures to **take** back *control* of their neighborhood - from new fences to strategic lighting - and don’t find the cameras intrusive.

(blogspot_outfile102; MGT, 2005)

The negative subcategory of the “power” tag (“S7.1-”) co-occurs significantly more frequently with the “M” nodes in the journal articles. The tag is realised by tokens such as *subject to* (Concordance 5-9, line 162), *powerless(ness)* (lines 163–165) and *rely* (line 168). These

examples suggest that the journal articles examine the power structures both in particular contexts (e.g. *border control*, line 170) and theoretically (“the idea of the panopticon”, line 169; cf. the theme of “Classic Surveillance” in Section 4.4.1). The journal is concerned with the “powerless”, but the detailed discussion of social processes and their *interdependencies*¹⁴, may add to the higher frequency of “S7.1-” lexis (such as *enslaves* in line 166; *reliant* in line 169) in comparison to the blog posts.

[162]	to expect that one 's	movements	will *subject* *to* enduring scrutiny
[163]	case. They noted that social	positions	characterized by *powerlessness* and by
[164]	-- adopted by those in *powerless*	positions	because they appear to reflect
[165]	heir lives. People in *powerless*	positions	may adopt paranoid beliefs then,
[166]	we are. Manning (1988: 155)	posits	that technology *enslaves* officers,
[167]	they *abided* *by* the rules	set	by a Federal Agency. For
[168]	thout personal experiences of the	districts	they have to *rely* on
[169]	that " the idea of the	panopticon	is completely archaic, " *reliant* as
[170]	discussed earlier in relation to	border	control. The equating of *uncontrolled*
[171]	1. The equating of *uncontrolled*	border	crossing with loss of sovereignty,

Concordance 5-9: Sample of (“M”, “S7.1-”) co-occurrences in the journal articles (10 out of 648)

As with the “S8” (“helping” and “hindering”) tags, the positive and negative subcategories of the “S7” (“power”) tag do not directly correspond to positive or negative evaluations of surveillant practices. To some extent, the examples reflect genre differences. In terms of mediated discourse analysis, the affordances and constraints of the text types lead to examples such as a reference to “social implications” (beyond the lines shown in this concordance). This expression is conditioned by the goals of research articles. By contrast, the content of the blog posts, shown in Section 5.4.1.2 to contain institutional language, appears to refer more to commercial and political roles.

¹⁴ *Interdependence* and *interdependencies* are realisation of the “S7.1-” tag in the S&S Corpus (seven occurrences).

5.4.2 Emotional aspects

In addition to the social aspects which I argued provide implicit evaluations of surveillant practices, this section is concerned with collocates related to emotion. The six main subcategories of the *USAS* tag related to emotional aspects (“E”) are listed in Table 5-8 and illustrated with examples.

Table 5-8: The first 20 distinct realisations of each “E” subcategory in the Surveillance Blogs Corpus (ordered alphabetically)

Tag	20 examples from the blog posts
“E1” (“General”)	affective, ambient, awed, compassion, curable, emotion, emotional, emotions, flush, flushing, gut, reaction, sensibilities, sentiment, sentiments, subjective, subjectivity, tone, undertones, witting
“E2” (“Liking”)	appreciate, cherish, dislike, enjoy, enjoys, favoured, for, hate, hatred, like, n’t, objections, popular, prefer, preferences, put, unpopular, up, was, with
“E3” (“Calm/Violent/Angry”)	abuse, abused, at, attack, attacks, coming, cross, force, grappling, hit, hits, kerfuffle, punches, rests, spite, threaten, threats, torture, tortured, torturing
“E4” (“Happy/sad”)	beam, celebrating, content, desperate, dissatisfaction, embarrassing, frustration, gratifying, grave, graves, grim, happiness, happy, laugh, pleasure, satisfaction, shattered, suffer, tragedies, tragic
“E5” (“Fear/bravery/shock”)	afraid, alarmed, chilling, cowardly, dare, fear, feared, fearing, fears, frightening, horrifying, horror, intimidation, menace, phased, scare, shocked, startling, terrifying, terror
“E6” (“Worry/concern/confident”)	apprehensions, care, concern, concerned, concerns, confidence, disconcerting, disturbing, faith, mind, neural, ordeal, qualms, trouble, trust, uneasy, worried, worries, worrisome, worry

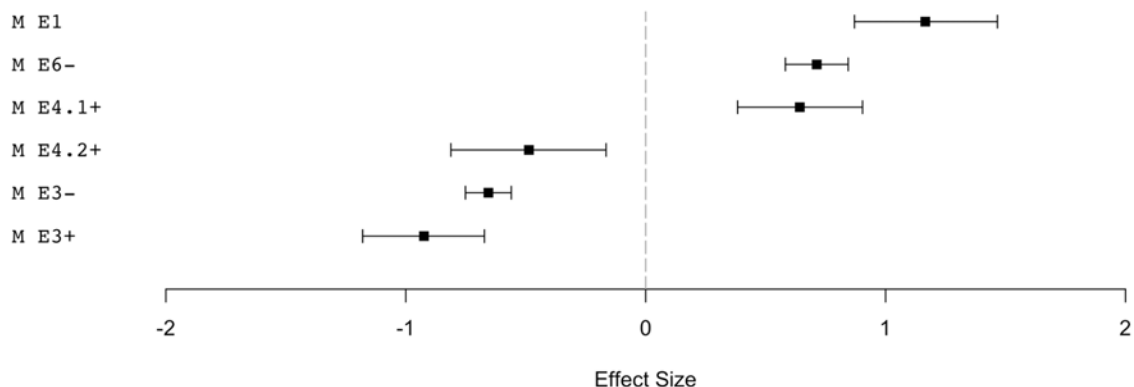


Figure 5-15: Co-occurrence comparison of the “M” semantic tag and “E” subcategories

Figure 5-15 shows the results of comparing the co-occurrences of “M” nodes with “E” subcategories across the corpora. Although the differences are not as stark as for the comparisons in Section 5.3 and even the social tags in 5.4.1, the results still point to distinct patterns. The bloggers have more freedom in expressing personal opinions about the surveillant landscape, whereas the journal articles aim to provide a more theoretical reflection of society’s responses to surveillance. The fact that surveillance may be only one topic among others, like news, marketing and economics, in the blog posts further differentiates the co-occurrence patterns.

Both the positive and negative subcategories of the “E3” tag co-occur significantly more frequently with “M” nodes in the blog posts than in the journal. Concordances 5-11 and 5-12 provide examples of each subcategory. The positive “Calm” tag is most frequently realised by *rest* (107 co-occurrences with “M” tags, e.g. Concordance 5-10, lines 54, 55, 57, 59), *peace* (lines 50–53)¹⁵ and *patient*. The occurrences of *peace* tend to appear in relation to movements, e.g. “policy, grassroots leaders and peace activists” (line 50), “peace and anti-intervention movements” (line 52) and organizations promoting peace (the “Peace Prize” in line 51 recognises the work of a media analysis website). Lines 58 and 59 are from a different, more commercial context in which a company advertises surveillance equipment. Example (34) shows the expanded context of line 59. It suggests that a “high quality burglar alarm” ensures that home is a “secure place”. So here, the “E3+” tag presents a surveillant landscape in a positive light.

¹⁵ The collocate **peace** is cut off at the beginning of line 53.

(34) Your home and your family are the most important; your home must be a **secure** **place** where your family can rest and relax. With all the criminal activity anywhere you can never take risks in protecting your home. [...] You need high quality burglar alarm system with 24 hours a day and 7 days a week monitoring capabilities.

(blogspot_outfile157; “Safety And Security In St. Louis/St. Charles,” 2013)

[50]	are directly affected by U.S.	foreign	policy, grassroots leaders and <i>*peace*</i>
[51]	was awarded the Gandhi Foundation International	<i>*Peace*</i>	Prize. ") The Medium and
[52]	the <i>*peace*</i> and anti-intervention	movements	. We stand in opposition to
[53]	nd anti-intervention movements. We	stand	in opposition to all forms
[54]	now, at <i>*rest*</i> and in	motion	. And then any animals that
[55]	them at <i>*rest*</i> and in	motion	. It would be rather inconsistent
[56]	atiently* crafted database of Wi-Fi	access	points, hotspots and cell towers,
[57]	the heavy lifting in the	slide	algorithms. The <i>*rest*</i> get passing
[58]	also allow you to enter	settings	like mood lighting for <i>*relaxation*</i> ,
[59]	home must be a secure	place	where your family can <i>*rest*</i>

Concordance 5-10: Sample of “M” nodes co-occurring with “E3+” (“Calm”) in the blog posts (10 of 388)

[57]	<i>*attack*</i> on Syrian territory, was	putting	new concerns on an already
[58]	of choosing the time and	venue	for their <i>*attacks*</i> , but normally
[59]	some, that the Pentagon had	crossed	the line from legitimate <i>*force*</i>
[60]	<i>*threatening*</i> is required to be	removed	from the TALON system in
[61]	ious incidents possibly linked to	foreign	terrorist <i>*threats*</i> to DoD resources,
[62]	, noting some terrorist <i>*attacks*</i>	where	pre-attack surveillance (by the terro
[63]	ce* protection <i>*threats*</i> would be	sent	to the FBI until a
[64]	<i>*threat*</i> . The group has numerous	front	groups, which dissolve and rename
[65]	utally* murdered Ayotunde Obanubi.	This	African student was <i>*attacked*</i> on
[66]	student was <i>*attacked*</i> on the	steps	of his college in east

Concordance 5-11: Sample of “M” nodes co-occurring with “E3-” (“Violent/Angry”) in the blog posts (10 of 2,302)

As a polar opposite of the peace-related lexis, the “E3-”, “Violent/Angry” tag is most often realised by forms of *attack** (539 co-occurrences with “M” tags in total; see e.g. Concordance 5-11 line 57, 58, 62, 65, 66), *threat** (407 co-occurrences in total; e.g. lines 60, 61, 63, 64) and *force* (143 co-occurrences; e.g. line 59, cut off at the beginning of line 64). These instances refer to situations of conflict or crime. Surveillance is often part of official investigations, but, as suggested in line 62, can also be used by terrorists as part of their preparation of attacks. Overall, in the Surveillance Blog Corpus, the “E3-” seems to refer to acts of violence rather than actual emotional responses to surveillant landscapes. Notably, the negative tag is almost six times more frequent than the positive subcategory. This appears to be

another example of the news value of negativity potentially playing a role in blog posts, in a similar way as noted in Section 5.3.2.2.

The “E4.2+” tag, which is also a significantly more frequent collocate of “M” nodes in the blog posts, is labelled “Content” and is realised by words like *pleasure* and *satisfaction*, but also *pride* and *complacency* in the Surveillance Blog Corpus. Example (35) illustrates the use of a “Content” collocate, *thankfully*, in a more casual tone than might be expected in the journal articles. Example (36) exhibits a similar casual, mocking tone.

- (35) One of the weirder tribes to have formed over the past decade is the hard core Republican supporter tribe. *Thankfully* **this** delusional and violent tribe seems to be a shrinking one, albeit one still capable of creating a lot of chaos.

(blogspot_outfile298; Big Gav, 2007)

- (36) Transnational Corporate Warlords are selling the same “military secrets” to our enemies. Just read the web sites of any top military contractor **where** they *proudly* proclaim their international business in military devices.

(blogspot_outfile349; “Who Really Owns and Controls the Military-Industrial Complex and What Are They Doing?,” 2016)]

The emotion tags on the side of the S&S Corpus relate to the journal’s role as a platform for conceptualising surveillance. Concordance 5-12 illustrates that the “general” emotion collocates include words like *subjectivity*, the most frequent realisation of the tag in the S&S Corpus (61 co-occurrences with “M” tags; e.g. lines 281, 283–285), *sentiment* (e.g. line 289) and *affective* (lines 287–288). These collocates do not represent a particular emotion but rather function as a metalanguage of emotions. So, this finding reflects the purpose of articles in the S&S Corpus to reflect on responses to surveillant landscapes in the real world more than to express the authors’ personal views. The only example in this concordance where the writer describes their own feelings is line 289. By contrast, the other lines reflect the role of the journal in analysing social structures (e.g. “social shaping of subjectivity”, cut off at beginning of line

283; “social and contextual change”, cut off at line 285) in relation to surveillance and its links to subjectivity and affect.

[281]	it is the product of	exterior	relations and of interior *subjectivity*.
[282]	the level of *ambient* sound.	This	sound level would determine the
[283]	shaping of *subjectivity*, and	this	reluctance inclines them to black-box
[284]	sted exploring the more lasting	intersections	between surveillance, *subjectivity*, and
[285]	ial and contextual change while	holding	the *subjectivity* of the individual
[286]	ediate, *visceral* contact with	ground	troops and so end up
[287]	embodied infrastructure of care,	foregrounds	the *affective*, embodied and gendered
[288]	fective* solidarities of ethnic,	national	and/or religious belonging and/or more
[289]	studies. I wish to echo	this	*sentiment* in suggesting more discussion
[290]	n 2015). Highly *emotive* cases	where	rapid warranted data access contributed

Concordance 5-12: Sample of “M” nodes co-occurring with “E1” (“Emotion: General”) in the journal articles (10 of 290)

Co-occurrences of the “M” tag with the “E6-” (“Worry”) tag are significantly more frequent in the journal and point to similar observations as the results for “E1”. In Example (37), the “Worry” tag is realised by *tensions*, which implies that surveillance is a contested social issue. This impression is reiterated with the phrase *ambiguity of dual conceptualisations* towards the end of the example. As for “E1”, the “Worry” tag may therefore reflect the focus of the journal on scrutinising the complexities of surveillance as a social concept rather than expressing actual distress. Additionally, in a similar way to “the territory of resistance” in Example (31), Section 5.4.1.2, the “M” tag (*surrounding*) here indexes a topic rather than a place.

(37) The complexities and *tensions* **surrounding** the surveillance of children in relation to food practices in residential care emphasise the ambiguity of the dual conceptualisations of children in need of both protection and control.
(2010-07-3_4-07; McIntosh, Punch, Dorrer, & Emond, 2010)

Unlike the “E4.2+” (“Content”) tag, the “E4.1+” (“Happy”) tag is found on the side of the journal. These categories appear semantically rather similar, which may explain why they do not exhibit large differences between the corpora. The arguably stronger “E4” category, “E4.1+” is realised by words like *amusing*, *playful* and *smiling*. In the S&S Corpus, these

emotion words tend to interpret and describe affective actions or situations in surveillant landscapes, as illustrated in Example (38), rather than evaluate them unlike e.g. *thankfully* in (35).

- (38) A growing number of transport operators are now supplementing ‘detached’ infrastructures such as CCTV and security personnel, with more ‘affective’ infrastructures of social control such as friendly *smiling* **platform** staff and welcoming station concierges at information desks.

(2013-11-3-09; Negishi, 2013)

5.4.3 Interactive surveillant landscapes

In Section 5.4 overall, the social dimension Jones’s (2017) surveillant landscape framework has guided the analysis of social and emotional aspects. Although no semantic tag can perfectly capture an interaction order, the “S” and “E” tag collocates have provided starting points for the analysis. Section 5.4.1 covered patterns related to (i) obligations and rights, (ii) benefits and problems of surveillant landscapes and (iii) power. The co-occurrence patterns revealed complex discussions surrounding questions of freedom, who is allowed to “walk free” or to publish texts or how people waive their privacy rights in everyday actions. Some examples criticise people in power (or the abuse of it) in relation to surveillance measures. Other examples have depicted surveillance as empowering, allowing residents to “take back control” of their neighbourhood, in a similar manner to the “empowerment” theme (Mehrabov, 2015) discussed in Section 4.4.1. A challenge for a corpus linguistic approach is that the wider context may reveal sentiments that are at odds with or directly criticise the statements in the concordance lines, which could be quoted from other sources.

The comparison of the “E” tag revealed some emotional aspects, such as the promotion of a home surveillance system for creating a *secure* place. Overall, however, these “emotional”

collocates rather pointed to the state of peace (or the lack of it) and surveillance in relation to violence and terrorist attacks in the blog posts. By contrast, in the journal articles, sentiments surrounding surveillance are discussed more frequently from a general, analytical level.

5.5 Conclusions

This chapter has conceptualised the relationship between surveillance discourse and place with a corpus linguistic approach to the notion of surveillant landscapes. I have argued that a number of parallels can be established between mediated discourse analysis – the basis of the framework of surveillant landscapes – and a corpus linguistic approach. The most important parallel is that both traditions focus on language in its social context.

The surveillant landscape framework was originally developed with a focus on semiotic and material elements in communicative situations. My analysis of the Bullring shopping centre has shown that mediated discourse analysis and corpus methods can be usefully combined to study both material elements and textual representations of a surveillant landscape. I have argued that the approaches to meaning-making in both traditions are based on co-occurrence patterns – whether words are co-occurring with each other or with other semiotic elements. In this chapter, the principles of mediated discourse analysis have provided a structure for the overall analysis and aided the qualitative study of meaning in specific contexts. The corpus linguistic approach has then taken the analysis to a larger scale as was done for the Surveillance Blog Corpus.

In analysing the example of the Bullring shopping centre, the chapter has identified a range of analogue and digital ways in which the centre's surveillant landscape is multimodally constructed (RQ 2-1). The discourses in place included surveillance signs (for CCTV and WiFi tracking), an anti-terrorism announcement and embodied surveillance by shop assistants and

armed police. In this landscape, digital surveillance was found to occur particularly via visitors' phones. The traces that this surveillance leaves behind, called historical bodies, have the potential to psychologically affect visitors' shopping habits, and to affect the centre's marketing strategies for individuals. Overall, the interaction orders suggest that visitors have little control in this surveillant landscape. This impression is reinforced by the textual patterns found in the privacy policy. An exception to this trend is an empowering feature of the centre's app that playfully encourages visitors to engage in surveillant activity by taking photos of fashion styles.

This chapter has argued that the interaction order is an aspect of the surveillant landscape that is of particular relevance to a corpus linguistic approach, but is also especially useful for working with the medium of blog posts. By writing public blog posts about surveillance, the bloggers inherently interact not only with their readers but also with their surroundings. Salient patterns in the blog posts, compared with the journal articles, have pointed to surveillance in relation to conflicts, politics, crime and privacy invasions of ordinary citizens, but also economics and commercial interests. The analysis has shown that the Surveillance Blog Corpus contains a large variety of attitudes to surveillance. In relation to interaction orders, discussions of freedom, rights, obligations and power have indicated the complex factors at play (RQ 2-2). In addition, technology-focused blog posts tend to show the positive capabilities of surveillance measures, at times in a commercial capacity. Terrorist attacks and the evaluation of political intervention are another important topic. By contrast, the academic articles tend to discuss the concept of surveillance more in terms of its connection with social structures and the tensions this creates. These differences relate to the affordances and constraints of the media, following the terminology of mediated discourse analysis. So, the two corpora represent registers with different social functions.

The semantic tags used in the second part of this chapter have facilitated the analysis of a large number of patterns related to surveillant landscapes. Accordingly, I was able to process a lot more data than by searching a corpus for particular words, not to mention a manual analysis as usually carried out by mediated discourse analysis. This larger scale analysis has simultaneously raised new challenges. Due to the heterogeneity and size of the Surveillance Blog Corpus it is difficult to understand the context of each individual blog post. Motivations for the blogs appear to range from personal interest to the promotion of ideological or political views, commercial services and products. Intertextuality plays an important role, not only for linking, but also in terms of incorporating text from sources like newspapers and Wikipedia. Despite taking a large-scale approach, the chapter has attempted not to lose sight of meaning-making patterns in context of examples with the guidance of the surveillant landscape framework. So, this chapter has contributed to the overall aim of the thesis by showing that surveillance discourse in the blogosphere is related to a range of surveillant landscapes which set up different social relationships.

6 Time: The diachronic representation of surveillance in *The Times*

6.1 Introduction

Meaning-making is crucial to my approach to surveillance discourse (see Chapter 2). This chapter focuses on the discourse coordinate of time. Meaning is not stable. It shifts over time as new paraphrases are added to the discourse (see Teubert, 2010). This notion of changing meaning over time fits in well with the theoretical model of ‘liquid surveillance’ (Bauman & Lyon, 2013; Lyon, 2010) that was introduced in Chapter 2. Accordingly, in the past decade, surveillance theory has accepted a more fluid character of surveillance as a concept which incorporates numerous factors unlike the clear hierarchy of powers imagined in more traditional accounts of surveillance. At the same time, liquid surveillance

[...] evokes the flows of data that are now crucial to surveillance as well as to the “time-sensitivity” of surveillance “truths” that mutate as more data come in (producing Kafkaesque consequences for some at the sharp end). (Lyon, 2010, p. 325)

Like meaning in discourse, these “surveillance truths” change as new information appears. Therefore, I argue in this chapter that by analysing the shifting representation of surveillance in news discourse over time, we can get a handle on how surveillance evolves as a concept. Whereas previous chapters have hinted at temporal factors in the discursive representation, for example across the journal volumes, neither the design of the S&S Corpus nor of the Surveillance Blog Corpus is suitable for a diachronic analysis. In the present chapter, the diachronic focus is supported by the temporal structure of the Times Digital Archive (TDA).

By investigating the third and final discourse domain of the thesis – news – this chapter moves from texts with comparatively limited audiences, as discussed in Chapters 4 (research articles in the *Surveillance & Society* journal) and 5 (blog posts), to a form of mass communication in *The Times* newspaper. As such, this chapter examines the most widely disseminated section of public discourse of this thesis. The chapter aims to address the overall question “how is surveillance discursively represented in *The Times* from 1986 to 2008, and how does this representation change across the period?”, which is split into four more specific subquestions. The time period of the analysis is restricted to the years 1986–2008 (the TDA198–2008 Corpus), because that period (i) saw a rise in the frequency of *surveillance* (see Section 6.2) and (ii) represents the highest OCR quality available in the local version of the TDA (see Section 3.2.3).

The TDA1986–2008 is a much more general corpus than the S&S Corpus and Surveillance Blog Corpus analysed in Chapters 4 and 5, which are both specialised on surveillance discourse. In order to find an entry point to the surveillance discourse in *The Times*, I chose the nodes *surveillance*, *privacy* and *CCTV* for the co-occurrence comparisons that address the first two RQs of this chapter: “How do long-term co-occurrence patterns in the surveillance discourse of *The Times* develop from 1986–2008?” (RQ 3-1) and “Which locally salient patterns are associated with relative frequency peaks in the surveillance discourse of *The Times* from 1986–2008?” (RQ 3-2). All three nodes have featured in the discussion of previous chapters, for example as KKWs of the S&S Corpus. *Surveillance* is considered a general representative of the discourse. Chapters 4 and 5 have shown that it is, unsurprisingly, at the core of much surveillance discourse, but does not cover all of it. It is therefore useful to include additional nodes, not least because *surveillance* can have negative connotations (see Section 5.4.1.3; Lyon, 2007, p. 15; also see Barnard-Wills, 2012, p. 87). *Privacy* is chosen as a concept

that is often associated with surveillance. The discussion of the patterns shows that this relationship is not straightforward (see Section 6.3.2; also Ball & Haggerty, 2005, p. 133). *CCTV* represents a prototypical example of a surveillance technology, especially before other technologies for mass surveillance raised public awareness with Snowden's release of NSA documents in 2013 (cf. Section 4.4.2.3).

In order to investigate long-term trends in the discourse (RQ 3-1), co-occurrences of the three nodes are compared across the years 1986, 1997 and 2008, i.e. the beginning, middle and end of the corpus. This top-down segmentation (see Section 2.6.2; Marchi, 2018, p. 179) of the corpus allows for an initial overview of the discourse in Section 6.2. In response to RQ 3-2, the analysis in Section 6.3 takes a more bottom-up approach. Subcorpora are selected based on peaks in the monthly relative frequency. Section 6.3 argues that relative frequency peaks provide a good place to find locally salient patterns in the surveillance discourse of *The Times*, as they point to particular events and news stories related to surveillance. The findings of this analysis complement that of Section 6.2 by supporting the longer-term trends and adding insights on specific events. The findings from Sections 6.2 and 6.3 both suggest that at different points in time from 1986–2008, patterns in the surveillance discourse are more associated with particular newspaper sections. RQ 3-3 (“How does the development of the surveillance discourse relate to newspaper sections in which the nodes occur?”) addresses this indication by tracing the frequencies of the nodes across the sections of the TDA online interface (GALE, 2018; see Section 6.4). This analysis of where patterns occur in a newspaper helps to contextualise the co-occurrence profiles, as the textual location of words and the placement of larger units such as entire articles contribute to the discursive meaning (Gupta, 2015; Mahlberg, 2007c; O'Donnell et al., 2012). The final part of the analysis moves from individual collocates to an investigation of the effect that an external event, the Identity Cards Act 2006, has on the

surveillance discourse (Section 6.5). This last analysis stage addresses RQ 3-4 (“How is the debate about the Identity Cards Act 2006 reflected in the surveillance discourse of *The Times*?”). Section 6.6 concludes the chapter.

6.2 Long-term co-occurrence patterns

This section presents the analysis of long-term co-occurrence patterns in the TDA1986–2008. Within the publication history of *The Times*, the coverage of *surveillance* appears to have first dropped, but then steadily increased. Figure 6-1 presents data in the period of *The Times* that is currently available via the online interface (Gale Cengage, n.d.), 1785–2013, on the number of documents mentioning *surveillance*.¹ The upper panel gives the percentage, the lower panel the raw number of documents. The percentage of documents is generally the one that would be considered more reliable because it represents normalised data. Yet, we can see from the raw number that the number of documents containing *surveillance* does not experience as sharp a drop from 1985 onwards as the percentages suggest. Instead, the total number of documents actually increase dramatically (from approximately 74,566 in 1985 to 129,760 in 1986). This may reflect the different underlying OCR dataset beginning in 1986 (see Section 3.2.3.2), or changes to the actual newspaper. Overall, Figure 6-1 indicates that the coverage of *surveillance* expands over the period of the TDA1986–2008, which suggests that the meaning-making patterns may also change across this period.

¹ The tool providing this data is called “Term Frequency”, but it provides only the number of documents rather than data on the frequency of *surveillance* itself.

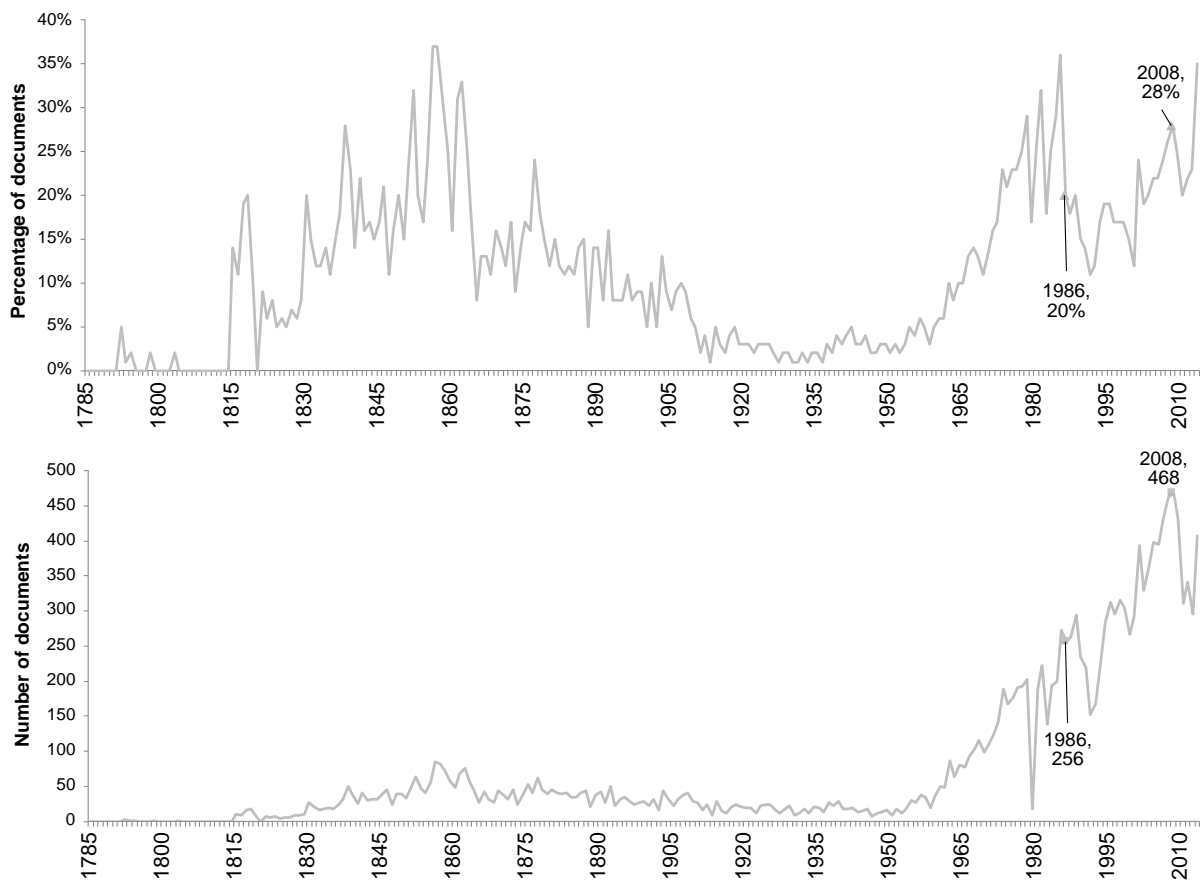


Figure 6-1: The number (upper panel) and percentage (lower panel) of documents in the TDA, 1985–2013, mentioning *surveillance*; data retrieved via the TDA “Term Frequency” feature

For an initial overview of long-term co-occurrence changes, co-occurrences of the node words are compared across the beginning, middle and final full years of the corpus as indicated in Figure 6-2, with three comparisons: (i) 1986 vs. 1997, (ii) 1997 vs. 2008 and, for the full length comparison, (iii) 1986 vs. 2008. The following subsections present the results of the three long-term comparisons for each of the node words with the help of *CorporaCoCo* results plots. The discussion starts with *surveillance* (6.2.1), which is given the most attention because of its central relevance to the discourse, followed by *privacy* (6.2.2) and *CCTV* (6.2.3).

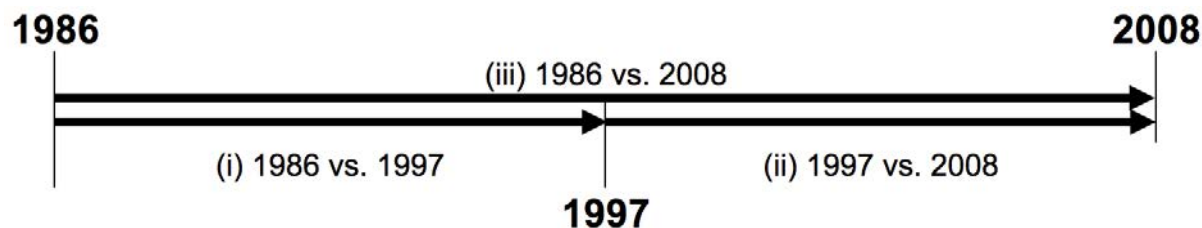


Figure 6-2: Long-term comparisons across the TDA1986–2008

Due to the nature of the comparisons featuring each year twice (e.g. comparing 1986 with both 1997 and 2008), the results across the two comparisons for each individual year overlap. Some collocates accordingly stay the same, irrespective of the reference corpus to which they are compared. However, the number of significantly different results differs across the comparisons; comparison (iii) of the first year against the final year of the corpus tends to yield the most results. A potential explanation is that the longer the gap on the temporal scale, the more co-occurrences tend to change. For each of the nodes, I discuss the yearly results chronologically for 1986, 1997 and, finally, 2008.

6.2.1 Long term comparisons for *surveillance*

Starting with the main node, *surveillance*, the results for the three long-term comparisons are shown in Figures 6-3 to 6-5. As the plots show, all three comparisons reveal significant co-occurrence differences for *surveillance*. All sets of results include collocates that appear in both years of the individual comparisons as well as collocates unique to either of the years. Most results are found for comparison (iii), which spans the full length of the corpus (see Figure 6-5).

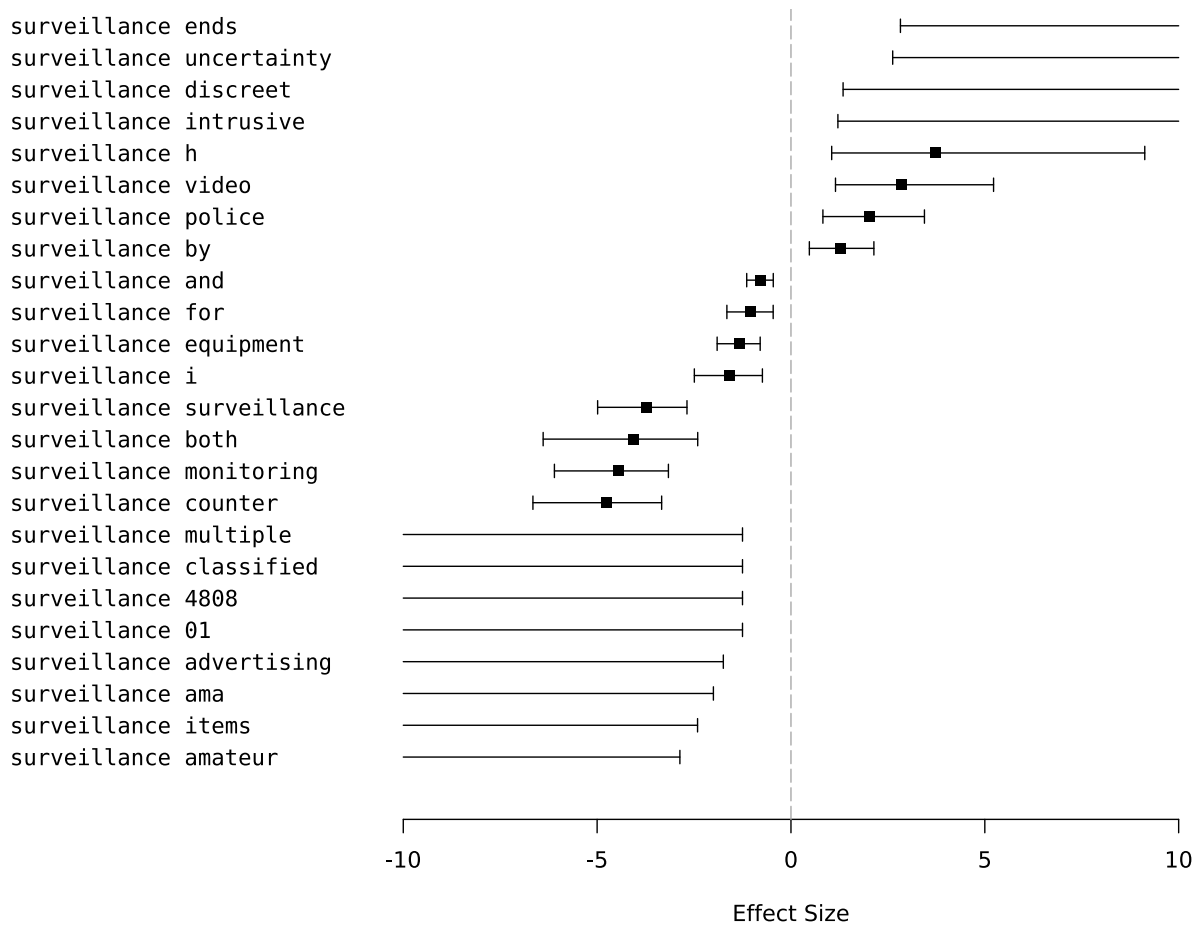


Figure 6-3: Co-occurrence comparison of *surveillance* in the year 1986 (left) vs. 1997 (right)²

² All *CorporaCoCo* plots and concordances in this chapter are based on the OCR-scanned local full-text copy of the TDA. Unlike the manually typed-up examples, they therefore contain OCR mistakes.

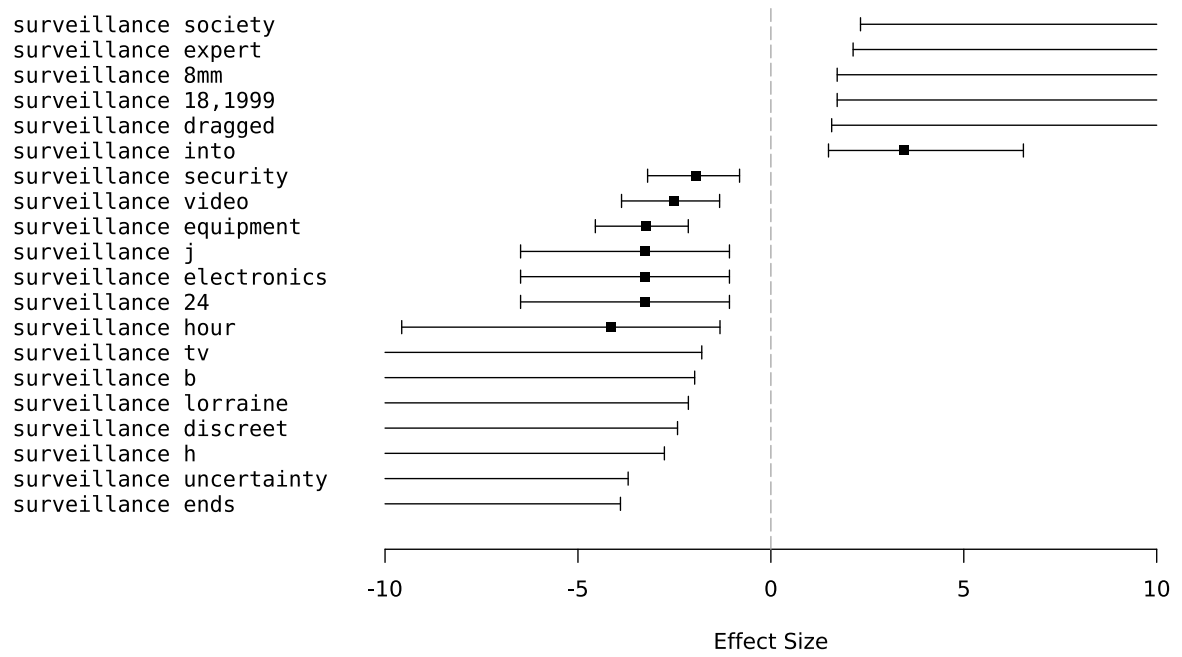


Figure 6-4: Co-occurrence comparison of *surveillance* in the year 1997 (left) vs. 2008 (right)

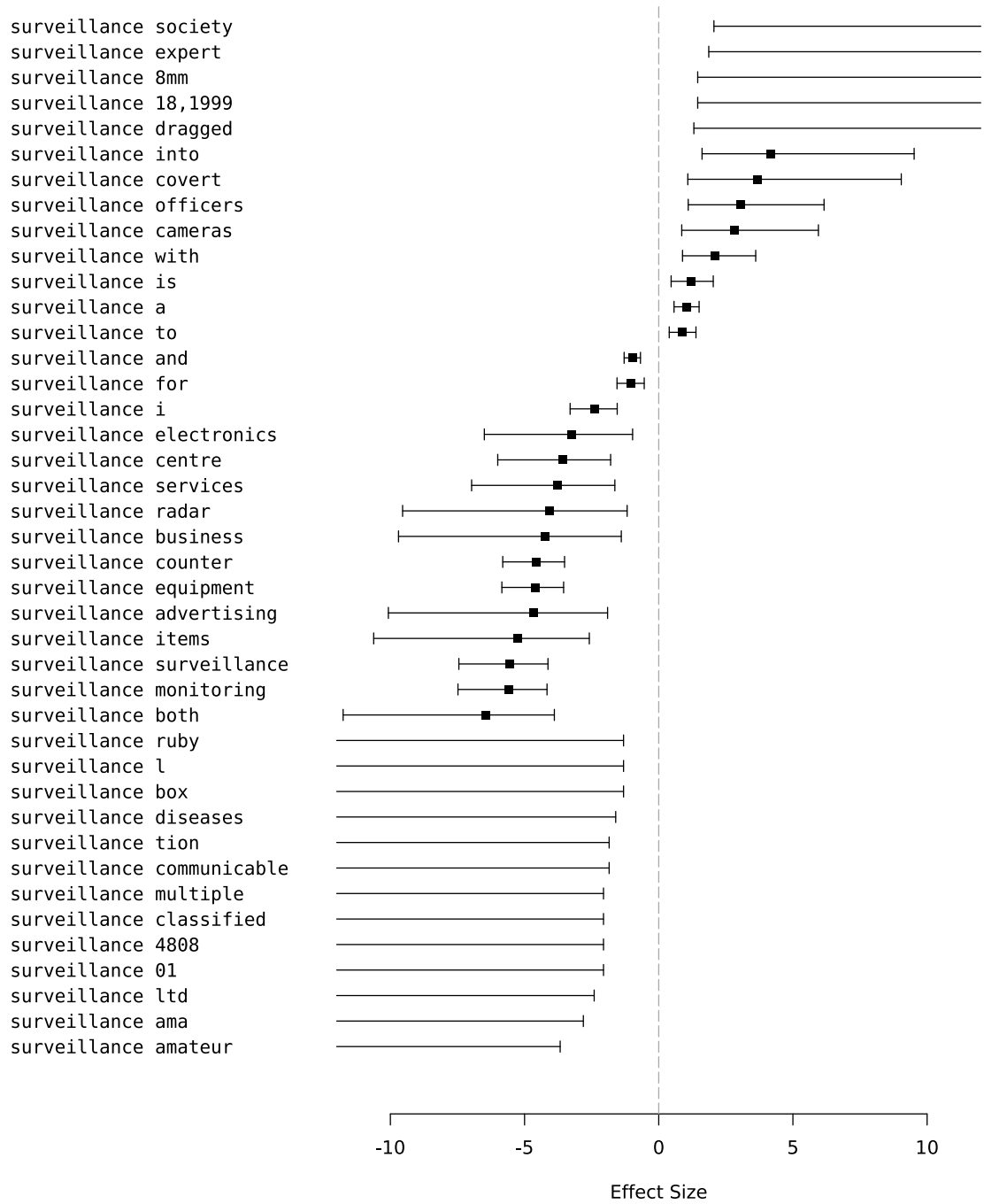


Figure 6-5: Co-occurrence comparison of *surveillance* in the year 1986 (left) vs. 2008 (right)

6.2.1.1 Surveillance collocates in 1986

The 1986 results point to early patterns in the development of surveillance discourse in the corpus and already indicate its importance as a cultural keyword. In both comparisons of 1986, many of the resulting collocates originate from the TDA category of “Multiple Classified

Advertising Items”. For example, the collocates *classified* and *advertising*, both found in Figures 6-3 and 6-5, relate directly to this section of the newspaper; they appear in section headings. Concordance lines of *surveillance* with the remaining 1986 collocates show that these majorly occur together in another specific advert. A search for “surveillance monitoring” in the TDA online interface for 1986 identifies the source as the advert shown in Figure 6-6 (“Ruby Electronics Ltd,” 1986), containing the collocates *monitoring*, *counter*, *surveillance*, *equipment both*, *amateur* and *professional*. The co-occurrences are likely so prominent, because this ad occurs in various *Times* issues of 1986.³



Figure 6-6: Advert containing *surveillance* and various 1986 collocates

In comparison to 2008, unique collocates in 1986 are *box*, from contact information in ads (*P.O. Box*) and *diseases*, which exclusively appears as part of the entity name *Communicable Diseases Surveillance Centre* in news articles. In her analysis of the cultural keyword *sustainable development* in news articles, Mahlberg (2007c, p. 200) identifies a group of concordance lines that appear in the names of organisations. Such occurrences in names of institutions, educational programmes or services are part of the “cultural” nature of cultural keywords, as in this case for *surveillance*.

³ Some versions of the ad differ slightly in typesetting with changes in font or hyphenation; e.g. in some instances *amateur* is hyphenated, leading to the result *ama* in the co-occurrence comparison, Figure 6-5.

6.2.1.2 Surveillance collocates in 1997

The co-occurrence patterns for the year 1997 also contain some advertising collocates. Both 1997 comparisons (Figures 6-3 and 6-4) have picked up *discreet*, *ends* and *uncertainty* as unique collocates of surveillance in 1997. Concordance lines of these collocates show that they appear together in the full text. A search in the TDA online interface identifies an ad by the company Lorraine Electronics as the source. Figure 6-7 illustrates the ad on one of the days that it was placed in *The Times*. In a similar way to the 1986 advert in Figure 6-6, the “Lorraine Electronics” (1997) advert promotes surveillance equipment. However, it appears to differ in its target audience. As suggested by the incorporated photo (see Figure 6-7), the advert mainly targets men, enabling them to check on their partner’s “disloyalty” with the help of “discreet surveillance”.



Figure 6-7: Advert containing co-occurrences of *CCTV* with *discreet*, *ends* and *uncertainty*

The 1997 collocates do not exclusively relate to advertising. For example, in comparison to 1986, *video* is more frequent in 1997 (see Figure 6-3). The collocates appear to reflect a growing awareness of surveillance evidence to be used for decision-making, not only in crime cases, but in the case of Example (1) even in professional sports. This article makes a strong case for *video surveillance* to be used in football matches in order to ensure the accountability of referees. Considering that the “VAR” (video assistant referee) system was only introduced by FIFA in

2018 (The International Football Association Board, 2018), the idea expressed in (1) appears rather futuristic.

- (1) Now does anybody really think that any match, in this era of over-protection, inter-player lawsuits and constant *video* **surveillance** could be so crazed, so wild, so vicious that one team needed to have four of its members actually removed from the supposed battlefield? [...] Everyone is up to speed with videotape and accepts its findings. If the referee gives a decision that a ball has not crossed the line and 20,000 voices scream that they are going to tear the ground apart based on that decision, he only has to halt play, and, via radiomike, ask for help.

(“The untouchables need electronic eye,” 1997)

Another example of evidence from video surveillance being portrayed as the solution to a problem is given in (2), in the context of crime. The detailed timestamps support the argument that this technology provides proof. Example (3) illustrates that the surveillance discourse at this point on the diachronic dimension engages in negotiation about the suitability of the surveillance measures, thereby contributing to meaning-making in Teubert’s (2010) sense. The instance of the co-occurrence pair in (3) problematises the use of the technology in the project described in (2), and the article further argues that the “tapes were of poor quality”. Time is an important factor: the two articles were published one day apart. The second article appears to be a reaction; at least to the project report that Example (2) is linked with, if not the article itself. These examples therefore provide evidence that the term *video surveillance* is being debated. The surveillance discourse develops as voices like these are added.

- (2) *VIDEO* **surveillance** by professor Southall and his team showed abuse by parents who exhibited warning signs the professionals had identified.

A three-month-old girl who had twice previously been admitted with breathing difficulties was observed with her mother from 12.24pm. Between 2.02 and 2.09 the mother slapped the infant’s head three times.

(“Video captures attack on baby,” 1997)

- (3) Colin Morley, a consultant paediatrician at Addenbrooke's Hospital in Cambridge, said yesterday that **video* surveillance* of abusers carried out by David Southall at hospitals in Stoke-on-Trent and London had a poor clinical record and exaggerated the true incidence of cruelty. [...] Dr Morley [...] said the tapes were of poor quality and were overinterpreted by doctors already suspicious of subjects.

(Henderson, 1997)

Another significantly more frequent 1997 collocate of *surveillance* compared to 1986 is *police*.

This collocate is no longer significant when 1997 is compared to 2008, suggesting that the co-occurrence rate of *police* is still high in 2008. Among the 1997 instances are references to a "Police Bill" proposed by the government. These include a mention in an article published on 21 January, which argues that the government may have to make "concessions" in its bill for protecting the rights of individuals captured by surveillance. The article also provides several examples of court cases using surveillance evidence in a problematic way, including *video surveillance*, which is one of the 1997 co-occurrence pairs from Figure 6-4, illustrated in Example (4). Example (5) is taken from a later paragraph of the same article. It is a quote from the Criminal Bar Association, apparently warning that evidence gathered via video surveillance and "bugging" should be safeguarded by standard procedures and warrants. The example demonstrates the tendency for privacy concerns to be closely linked to debates over surveillance powers.

- (4) In the private prosecution for murder [...], the teenager stabbed at a bus stop in southeast London, **video* surveillance* from inside the home of a suspect was shown in committal proceedings. The trial collapsed, however, and the tapes were not seen by a jury.

(Gibb, 1997)

- (5) "In all other areas where the state is empowered to interfere with the privacy of the individual, the sanction of a judge, a magistrate or the Secretary of State is required [...]"

(Gibb, 1997)

Home Secretary replies on bugging

From the Home Secretary

Sir, Dr Ian G. Bogle of the British Medical Association and others (letter, March 10) express continuing concerns about the Police Bill's provisions on intrusive surveillance.

We have tabled substantial amendments to the Bill in response to the concerns that were originally raised. If an operation is in a sensitive category, including any that are likely to intrude on medical confidentiality, the police will normally be required to obtain the approval of a commissioner — who will be a serving or retired High Court judge — before that operation begins.

Figure 6-8: Beginning of the response by the Home Secretary to previous letters to the editor

Figure 6-8 shows a response by the Home Secretary to previous letters to the editor, published in *The Times* on 12 March (Howard, 1997). It quotes concerns over “intrusive surveillance” by the opponents of the Police Bill published in letters to the editor on 10 March. With this response, the Home Secretary apparently attempts to regain the public’s trust by promising “substantial amendments”. So, the discussion of the Police Bill is a prime example of negotiation in the discourse and *The Times* being an important platform for this.

6.2.1.3 *Surveillance* collocates in 2008

In both comparisons of 2008, the top unique collocate *society* is mostly realised by mentions of the compound *surveillance society*, illustrated in Concordance 6-1.⁴ The analysis of the blogs in Chapter 5 identified this bigram as part of the criticism of the US government, claiming citizens are being “tricked” into “a total surveillance society” (Section 5.4.1; blogspot_outfile316; Stephens, 2015). The patterns surrounding *surveillance society* in the

⁴ *CorporaCoCo* marks the boundary between the span used for the co-occurrence comparison – a span of five in this thesis – and the extended context with three dashes (---).

2008 subcorpus of *The Times* have a similarly negative connotation. A salient theme in Concordance 6-1 portrays Britain as either being (at risk of) “turning into” a surveillance society (line 13; also, *becoming* in lines 26, 28, 29). This idea is also expressed metaphorically (e.g. “on the road to”: line 8; “another step towards”: 9), with the phrase “sleepwalking into” occurring several times (lines 11, 19, 27).

```
[ 1] click on timesonlinex.co.uk/nioysetrap Surveillance society of council officials into ---
[ 2] --- society." The report, entitled A Surveillance Society?, said that it accepted ---
[ 3] ou --- shudder. The Lives of Others. Surveillance society. Stasi Britain. All that. ---
[ 4] thousands of council officials into ... Surveillance society Terror law turns thousands ►
[ 5] r: --- One Man's Struggle against the Surveillance Society The Times THE JMiTIMES --- is
[ 6] ntrol and no --- remedy. We live in a surveillance society. A report this week --- for
[ 7] eaves the --- stalls. A satire on the surveillance society?A bit. IGT IGT --- MARILYN
[ 8] is another step "--- on the road to a surveillance society" and has warned the --- airpor
[ 9] e move --- was another step towards a surveillance society and warned BAA that --- it
[10] fantastic tool' vestigation into the surveillance society and is to. publish --- its
[11] re at --- risk of sleepwalking into a surveillance society, at least it will --- be a
[12] cer --- will be questioned about the "surveillance society" by the Home Affairs ---
[13] ars that the --- UK is turning into a surveillance society. "CCTV enjoys a lot --- of p
[14] there is a --- bit of comment on the surveillance society. "For your attention,"
[15] dress it for --- you?" If this is the surveillance society, give me more. Fly --- Away
[16] --- in Burnhamon-Sea, Somerset. The 'surveillance society' has led to many ---
[17] still --- be satisfied with what the "surveillance society" has achieved. Many innocent -
[18] A database --- and the growth of the "surveillance society". He said that he --- was
[19] that we --- are sleepwalki ing into a surveillance society. Hold\ ing large collections -
[20] y a few --- hours of freedom from the surveillance society. I refer, of course, --- to
[21] er said: "--- Our concern is with the surveillance society. Is this m A7r»*--- ii another
[22] --- writes). Villagers claim that the surveillance society of modern Britain's towns ---
[23] ople --- audio guide." cameras, as he surveillance society, rather been Davis," he ---
[24] lism (--- Rumours of a Hurricane) and surveillance-society satire (The Seymour Tapes). --
[25] as part --- of its inquiry into the "surveillance society". Shami Chakrabarti, the
[26] cerns --- that Britain was becoming a surveillance society. Shami Chakrabarti, the
[27] --- other day - "sleepwalking into a surveillance society". The Oxford academic and ---
[28] that the --- UK is slowly becoming a surveillance society, the committee has looked ---
[29] se we are --- in danger of becoming a surveillance society." The report, entitled A ---
[30] 4). This, --- is an extension of the "surveillance society" too far. I support --- the
[31] t's --- remember the benefits of the "surveillance society". We should draw satisfaction
[32] It said: "--- Our concern is with the surveillance society. Why do they need ---
```

Concordance 6-1: All 32 instances of *surveillance* co-occurring with *society* in the R1 position in the TDA, 2008

Another prevalent theme is that Britain is a *surveillance society* already, as apparent from the statement “We live in a surveillance society” (line 6) and suggested by the use of the definite article (lines 5, 7, 10, 12, 14-18, 20-22, 25, 30-32). These lines are examples of explanations and paraphrases of the term *surveillance society* that contribute to the ongoing negotiation of meaning in surveillance discourse. Some of the lines with definite articles go further by arguing that the *surveillance society* does not only exist, but its extent is expanding (“growth of”: line

18; “extension of”: 30). Line 15 is a curious case because the bigram occurs in an *if* clause: the extended context shows that this example from a column refers to a positive experience with the Royal Mail customer service who had called the writer to follow up on an incompletely addressed postcard. Other playful examples are the seemingly ironic use of “a few hours of freedom from the surveillance society” (line 20) from an article about election fraud, which criticises the lack of processes in place to check voters’ identities and the “satire on the surveillance society” (line 7).

The public sentiment that Britain is (becoming) a surveillance society expressed in Concordance 6-1 has also been discussed by surveillance studies scholars. Barnard-Wills (2012, p. 16) suggests that the term *surveillance society* was introduced to the media through a report for the Information Commissioner, but was already used by surveillance studies scholars before that:

The United Kingdom Information Commissioner has stated that he fears the UK is ‘sleepwalking into a surveillance society’ (BBC News 2004) and the term is now commonly used in the media, having spread through press releases and even finding its way into party political manifestos at the last election. The phrasing mirrors the title of the report commissioned by his office from the Surveillance Studies Network (as well as a 1989 article by Oscar Gandy and a 2001 book by David Lyon) [...]

The observation that the currency of the term increases across the 2000s is supported by the frequency development of the bigram in the TDA1986–2008 (see Figure 6-9). Accordingly, *surveillance society* first appears in the corpus in 1998, and gradually increases from single mentions to a raw frequency of 32 in both 2007 and 2008. The term is similarly featured in the blog data in Example (27), Section 5.4.1.2.

The quote by Barnard-Wills (2012) explains the popularity of the *sleepwalking* expression seen in Concordance 6-1 for the year 2008. He points out that the expression is not coined in 2008. The TDA data confirms this point, as the phrase already occurs in the concordances of previous years (2004, 2006, 2007), along with similar imagery evoked by the

expression *waking up to a surveillance society* in 2006 (Ford, 2006). In 2004, the year also mentioned by Barnard-Wills (2012) in relation to the Information Commissioner, four out of the six concordance lines of *surveillance society* contain that phrase. These instances mention a potential identity cards scheme as a major concern in the development of the *surveillance society* (see also Section 6.5). The metaphors of “sleepwalking” and “waking up” suggest a feeling of powerlessness in the face of an apparently unprecedented spread of surveillance across wide sections of society.

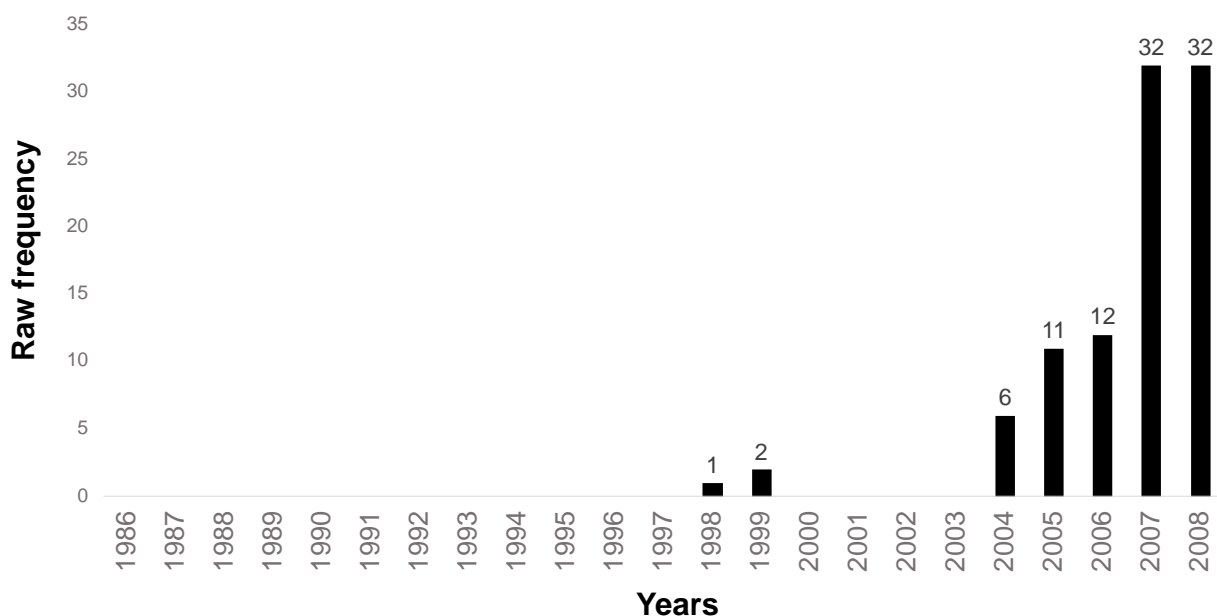


Figure 6-9: Raw frequencies of *surveillance society*

Cameras is significantly more frequent collocate in 2008 compared to 1986. The higher usage likely relates to the increase in the use of surveillance cameras across the corpus period (also seen in the increase in mentions of *CCTV*; see Section 6.3.3). With the increasing use of the technology it has become more widely used in public places, and, therefore, also covered in the press, as indicated in the 1997 examples.

[1] 4 Intelligence --- Company (14 Int), a covert **surveillance** unit, had not taken a --- May
[2] d placed --- under covert online and physical **surveillance**. Interviewers were coached for
[3] - about --- high-profile prisoners and covert **surveillance** in the prison system. What ---
[4] es that he had --- no knowledge of any covert **surveillance** in his branch. He said ---
[5] hed. Flat --- 386a was routinely "burgled" by **surveillance** teams who planted covert video
[6] , but they do --- allow one to *conduct* covert **surveillance** with impunity. Novv that I ---
[7] ils --- are routinely under covert electronic **surveillance**, security sources told The
[8] for his --- work in *conducting* covert aerial **surveillance** of Taleban positions, co-
[9] April 2007 --- and July 2008, covert directed **surveillance**, as authorised under the
[10] s the --- validity of her fundraising. Covert **surveillance**, once the stuff of John --- le
[11] authority --- in other agencies Human covert **surveillance** Undercover agents (informants
[12] aissance --- Regiment (SRR), expert in covert **surveillance**. However, the only other
[13] ussed the --- merits of various items. Covert **surveillance** produced photographic evidence
[14] oners are --- routinely the subject of covert **surveillance**. The revelation comes days
[15] pect, to --- investigate the extent of covert **surveillance** in Britain's prisons. Mr
[16] lities of directed surveillance - "the covert **surveillance** of a --- particular person, or
[17] nce of intercepts --- is that tapes or covert **surveillance** could be doctored. This is ---
[18] is all --- glamour, right? Stake-outs, covert **surveillance**, wealthy, charming paymasters
[19] been stripped --- of its control over covert **surveillance** teams in an attempt to ---
[20] al victim is safe - --- we would not rule out **surveillance**, we may use covert tactics ---
[21] rybody's goat. --- And the reason that covert **surveillance** annoys people is that we ---
[22] known to the --- group as Dawood. The covert **surveillance** and the intelligence from
[23] CCTV and other --- cameras. It is the covert **surveillance** that *has got everybody's goat*.
[24] es --- of directed surveillance - "the covert **surveillance** of a particular person, or ---
[25] hing for anti-*terror police* to use covert **surveillance**, but it has come to --- a
[26] ow alarmed --- His *officers* didn't use covert **surveillance** in investigations connected to
[27] Rangzieb --- Ahmed was tracked using covert **surveillance**. Listening devices had been

Concordance 6-2: All 27 instances of (*surveillance*, *cover*)

The 2008 collocate *covert* refers to secretive surveillance measures. The concordance lines for this co-occurrence pair (see Concordance 6-2) differ from the more abstract discussions in Concordance 6-1. Instead, here, the main themes are concerned with questions of agency. These relate to who is the subject of surveillance, such as prisoners (lines 3, 15), the “Taleban” (line 8) or “a particular person” (line 24) rather than the mass, and who is “conducting” it (see lines 6, 8). The agents include a “cover surveillance unit” (line 1), “teams” (lines 5, 19), “undercover agents” (line 11), “terror police” (line 25) and “officers” (line 26).

One question that appears to be of particular importance is which branch of the police can legitimately use covert surveillance. Example (6), which gives the extended context of line 25, distinguishes between covert surveillance being used in anti-terror investigations – where it is presumably acceptable – and for school catchment area checks. The implication is that ordinary people should not be the subject of covert surveillance. This sentiment resonates with the depiction of covert surveillance as a measure that “annoys people” (line 21; also see line 23 from the same article).

- (6) James Welch, legal director of the human rights pressure group Liberty, said: “It’s one thing for anti-terror police to use *covert* **surveillance**, but it has come to a pretty pass when it becomes the tool of the school catchment area police.”

(Bruxelles, 2008)

Most remaining 2008 collocates of *surveillance* relate to the film *8mm* that repeatedly appeared in the *Times* section “Today’s TV” with the description (or slight variations of it) shown in Example (7). The numbers *18,1999* (as the form is OCR-recognised and shown in Figure 6-5) refer to the age rating (18 years) and (1999) refers to the release year. In a similar way to the appearance in the company name in the 1986 ad, this mention of a “surveillance expert” in a film synopsis illustrates the cultural nature of the cultural keyword *surveillance*.

- (7) *8mm* (*18, 1999*) A **surveillance** *expert* is *dragged* into the seedy world of murder and deception when hired to investigate a suspicious movie. Starring Nicholas Cage

(Ed Potton, 2008)

6.2.2 Long term comparisons for *privacy*

In the previous section, co-occurrence patterns have indicated that surveillance discourse does at times involve concerns over privacy being compromised, as in Examples (4) and (5) about surveillance evidence used in court. This section directly focuses on the patterns of *privacy*, following the same set-up of yearly comparisons. Figures 6-10 to 6-12 present the significantly different co-occurrence pairs across the years.

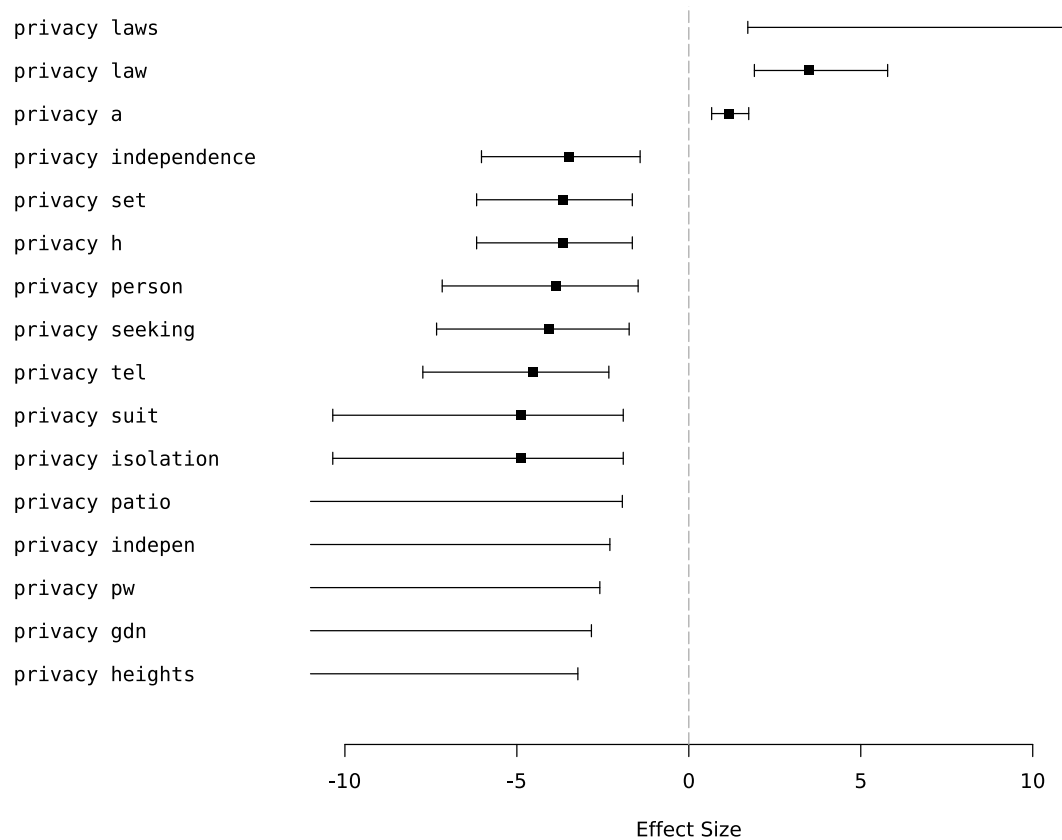


Figure 6-10: Co-occurrence comparison of *privacy* in the whole year 1986 (left) vs. 1997 (right)

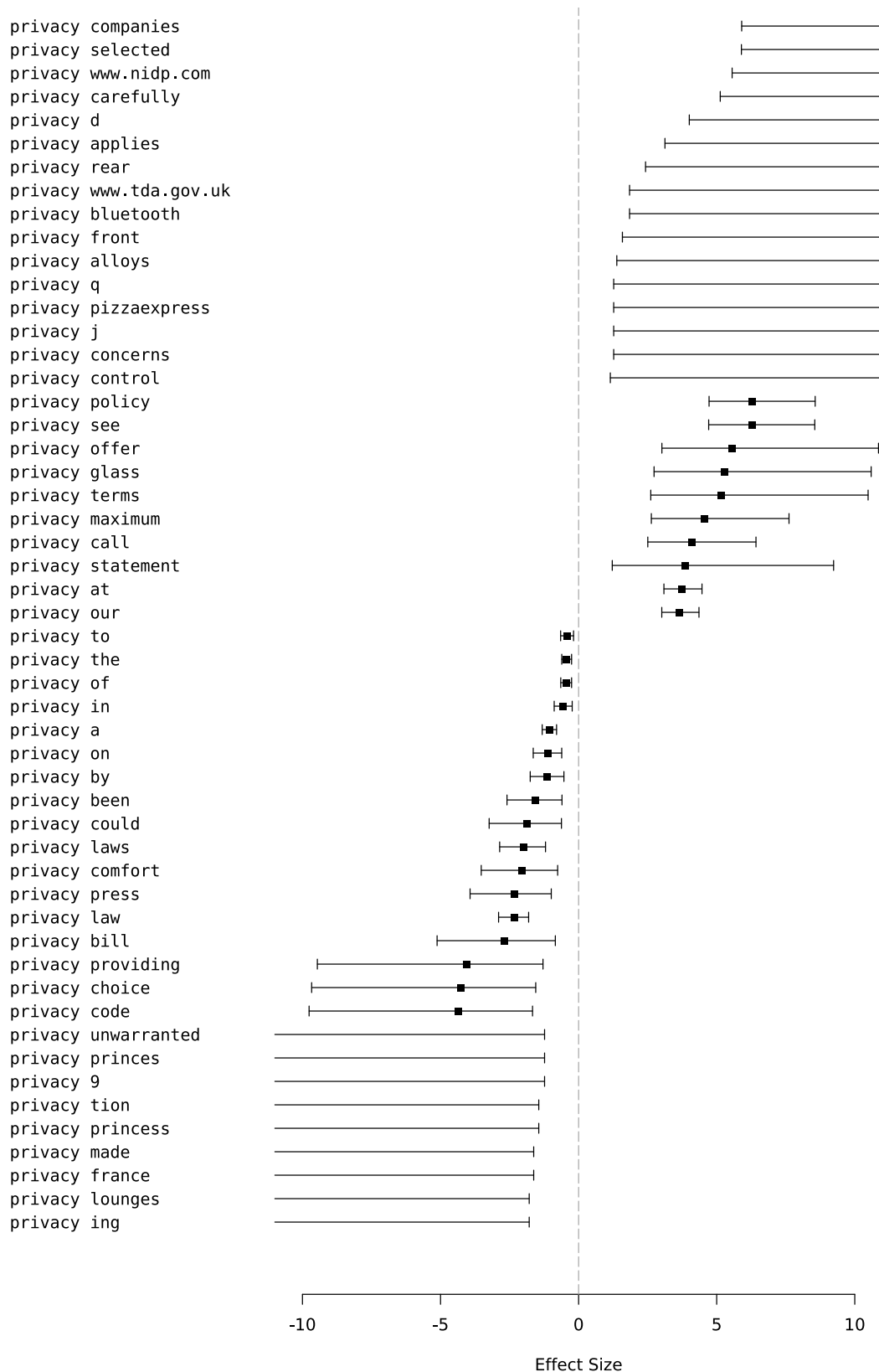


Figure 6-11: Co-occurrence comparison of *privacy* in the whole year 1997 (left) vs. 2008 (right)

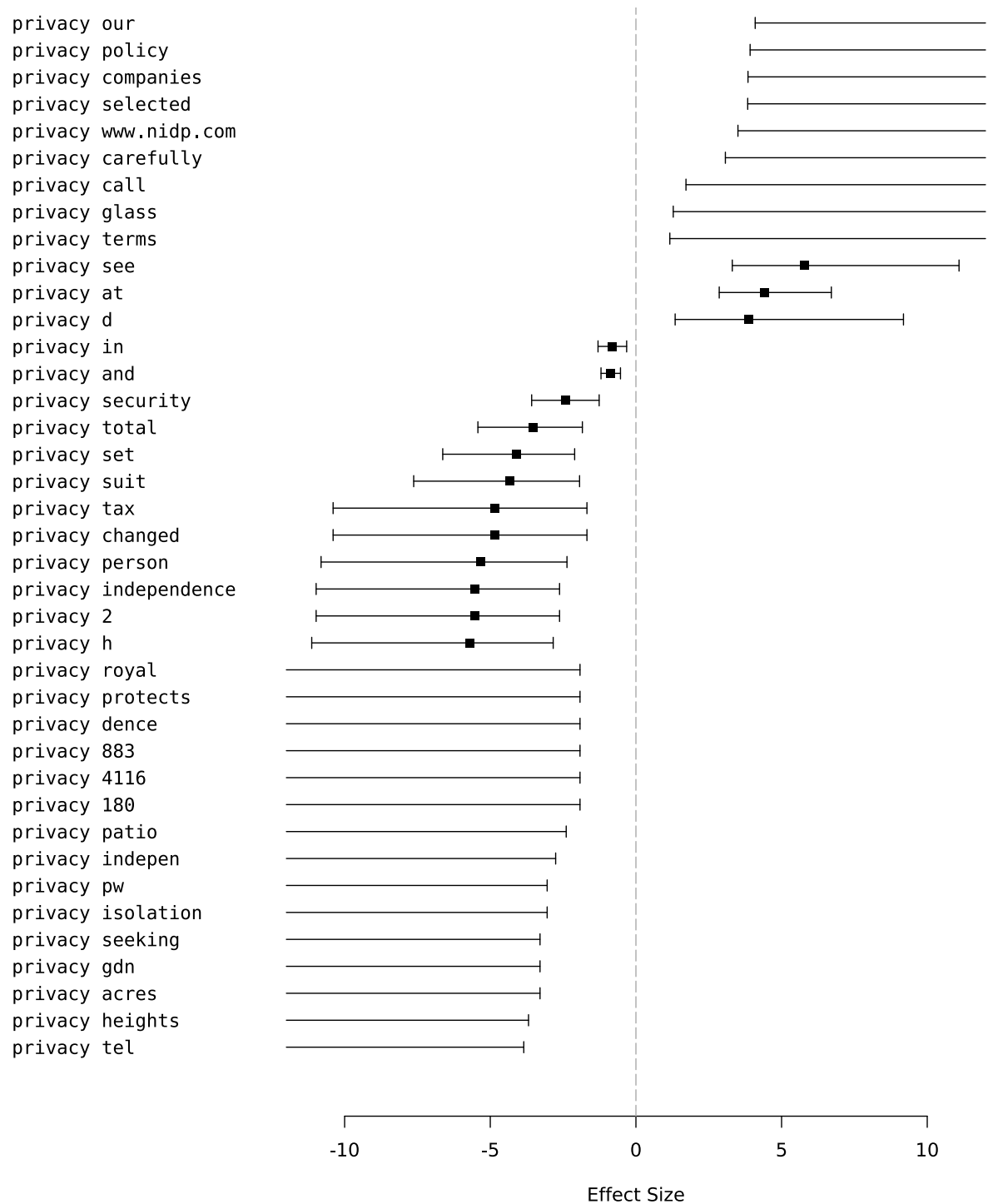


Figure 6-12: Co-occurrence comparison of *privacy* in the whole year 1986 (left) vs. 2008 (right)

6.2.2.1 *Privacy* collocates in 1986

Many of the *privacy* collocates found in 1986, compared to the other years, appear to relate to housing ads. For example, the ad in Figure 6-13 (“Rentals,” 1986) occurs in various *Times* issues and contains several of the 1986 collocates of *privacy* from Figures 6-10 and 6-12, including *patio*, *suit*, *seeking* and housing-specific abbreviations such as *gdn* (“garden”) and *pw* (“per week”). Abbreviations like these are conventional markers of classified advertising and housing ads in particular; in his study of US classified ads, Bruthiaux (1996, p. 95) identifies the highest percentage of abbreviations in ads of the categories “autos” and “apartments”.

EXCLUSIVE NEAR HAMPSTEAD
Highly Prestigious, for lux fur-
nished flat in quiet res area. 2
bed. 2 rec. k & b. Patio & Gdn.
Suit Person Seeking Privacy
£180 pw. Tel:01-883 4116

Figure 6-13: Example of *privacy* co-occurring with *suit* and *seeking*⁵

Further housing-related examples for *privacy* from 1986 are given in Concordance 6-3 for the co-occurrence pair (*privacy*, *isolation*). These patterns remind us that privacy is an asset to our daily lives and especially our homes, as King (2004) studies in depth in *Private Dwelling*. Whilst the ads may initially seem like “noise” in a study of newspaper discourse, they still reflect social values. Line 3 stands out from the pattern of housing ads; instead of praising the lack of isolation it mentions “feelings of isolation”. Example (8) provides the extended context,

⁵ This is an example of a TDA “document” that has been misallocated to a section. Although the text clearly appears to be advertising, it appears in the “News” section of the TDA online interface (and the XML categories would be equivalent). This is also an example where the quality of the scan itself seems rather poor, which can easily have been reflected in the quality of the OCRed text. Nevertheless, collocates like *suit* and *seeking* have been correctly identified.

originating from an article on mental health issues experienced by North Sea oil workers. Compared to the housing ad, this example describes almost the opposite situation: feelings of isolation occur due to the remote location and despite a lack of privacy as too many men live in a confined space. Both contexts suggest that privacy is important for a good living environment, even though it interacts with many other factors.

```
[1] floor layout. 2 dble bedrms. Privacy without *isolation*. £47.500.
[2] floor layout . 2 able bedrms. Privacy without *isolation*. £47.500.
[3] of job dissatisfaction , lack of privacy, feelings of *isolation* and
[4] bath & kitchen). 2 Acres. Total privacy without *isolation*.
[5] iath & kitchen). 2 Acres. Total privacy without *isolation*. £300,000
[6] & fliitchenj. Z.'Acres . Total privacy ; without
[7] bath & kitchen). 2 Apies. Total privacy YriHioiA *isolation*. £300000
```

Concordance 6-3: All 7 examples of *privacy* co-occurring with *isolation* in 1986

- (8) Many of the men complained of job dissatisfaction, lack of **privacy**, feelings of *isolation* and difficulties with relaxing with their families when they returned to shore after up to 21 days on the platforms.

(Davenport, 1986)

6.2.2.2 *Privacy* collocates in 1997

By contrast to the housing patterns of 1986, the 1997 collocates *law* and *laws* – dominating in Figure 6-10 and still present in Figure 6-11 – highlight a theme that is more closely related to the main news section of *The Times*. An article from April 1997 discusses the formation of a new commission, which would develop a “new code of practice on privacy and fair treatment” for the media, as illustrated in Example (9). By calling this exercise an “unhappy” one – and reflecting on that choice of words – the extract points to a general tension between protecting the privacy of individuals and the role of the press to inform the public. In particular, this example focuses on the need to expose injustice and illegal activities that those individuals may be involved in. The extract in (9) starts with the opening paragraph of the article, which takes a

rather pessimistic stance in the face of modern technologies including “surveillance cameras”, proclaiming that “privacy is dead”.

- (9) Privacy is dead. Camcorders, surveillance cameras, cash machines and credit cards record every aspect of our lives. [...] With no secrets left, how to protect the unprotectable? The new Broadcasting Standards Commission is under statutory obligation to try. Formed at the beginning of the month [...] the new commission has the unhappy task of drawing up a new code of practice on privacy and fair treatment. I say “unhappy” because it is impossible. All codes and **laws** to protect **privacy** run the risk of concealing injustice.

(Maddox, 1997)

- [1] g'^^^ SHH^WBM1^ fl^BSBBBBBBB^S^^^^ Princess photos 'will not --- lead to press curbs' NEW **privacy** **laws** were ruled out yesterday --- amid controver sy over the publication of pictures of Diana
- [2] the death of the Princess is likely to force consideration --- of the introduc tion of **privacy** . **laws** and the Commons National --- Heritage Select Committee is expected to hold a full inquiry
- [3] Bodkin It is a *tragic irony* that France,' the country --- with one of Europe's strictest **privacy** **laws**, is where Diana, Princess --- of Wales was killed, pursued by paparazzi. It is in
- [4] of the Princess's death. A further 76 per cent demanded --- tougher **laws** to protect the **privacy** of public figures. Last night, --- John Howard, the Prime Minister, gave a warning of the
- [5] Diana, Princess of Wales, is the *ideal model* around whom --- to construct a **law** of **privacy** . Her impassioned interview with Le --- Monde yesterday painted a heartrending picture of a Princess at
- [6] Foreign Secretary, yesterday issued a veiled warning against the media --- of the possibility of a **privacy** **law** in the wake of --- the Princess's death. Mr Cook, who was the first government
- [7] Birch Reynardson, Oxford county councillor Page 25 : '^ ;//; LETTERS .; ; " , ' . "' Princess of --- Wales; harassment and use of **privacy** **law**; MI5 activities; Scottish banknotes; --- in-cell tele vision ; metric rain Page 23 â " El Mundo
- [8] Paris is as inevitable as it is understandable. Let us --- not pretend, however, that any **privacy** **law** con ceived in the --- emotion-charged after math of Princess Diana's untimely death could
- [9] committee yesterday, a message was read from the Princess's brother --- Earl Spencer calling for a **privacy** **law** and an end to --- the "torture" of privacy invasion. Afterwards Sir David Eng lish
- [10] Murdoch warns against privacy law RUPERT MURDOCH urged the media --- to resist demands for a **privacy** **law** yesterday in the wake --- of the death of Diana, Princess of Wales. He said
- [11] newspapers after the Princess's death, meanwhile strongly resist ing the --- introduction of a **law** of **privacy**. They would hardly help their --- cause if they published Prince Harry's picture, if the Palace

Concordance 6-4: All 11 examples of *privacy* co-occurring with *law* or *laws* within a span of five and containing *princess* or *princess's* in an extended context (span of 15) in the 1997 subcorpus out of 188 total instances⁶

⁶ There are 52 instances of (*privacy*, *laws*) and 136 of (*privacy*, *law*). The extended context is marked off by dashes (---).

Other reasons for the conflicts of public interest and privacy may simply involve the prominence of the individual and questions of publishing details and images from their personal lives. As a case in point, the concordance lines of *privacy* with collocates from Figures 6-10 and 6-11 show that a particular person plays an important role in the 1997 debate on privacy law: Diana, Princess of Wales. Concordance 6-4 shows instances of *privacy* co-occurring with *law* or *laws* in 1997, filtered to those that also contain *princess* or *princess's* in an extended context of an additional ten words.

The concordance points to two main contexts of privacy law discussions in which the Princess is mentioned. These can be distinguished chronologically: the first context is the “controversy over the publication” (see line 1; from an article on 11 August) of photos that show her in intimate moments with her new partner. At this point in time a change in law seems unlikely (see line 1: “NEW privacy laws ruled out”). Line 5 also refers to this context. As Example (10) shows, the article actually questions whether the Princess “is the ideal model” for the argument of protecting an individual with privacy law. The example quotes further chunks from the same article (published on 28 August), which argues that the privacy of public figures fundamentally differs from the privacy of ordinary individuals, because it is “a privacy of a peculiarly public nature”.

- (10) I doubt whether Diana, *Princess* of Wales, is the ideal model around whom to construct a *law* of **privacy** [...] Neither the Princess nor the Minister are in serious need of protection, being well-versed in the ways of the media. Both are experienced in allowing information to be disseminated when it suits them [...] It is, I suppose, an invasion of **privacy**, but a **privacy** of a peculiarly public nature.

(Linklater, 1997)

The second context dominates the concordance, relating to the fatal car accident in which the Princess was involved on 31 August 1997. The death of the Princess is referred to repeatedly

(lines 2-4, 6, 8, 10, 11), thereby establishing a link between the incident and the discourse of privacy laws, although views are mixed.

On the one hand, “in the wake of the Princess’s death” (line 6), calls are made for strengthening the legal protection of privacy (lines 2, 4, 6, 9). The Foreign Secretary warns the media of the “possibility of a privacy law” (line 6), the “Princess’s brother” calls “for a privacy law and an end to the ‘torture’ of privacy invasion” (line 9) and Australian survey respondents “demanded tougher laws to protect the privacy of public figures” (line 4). On the other hand, newspapers are “strongly resisting the introduction of a law of privacy” (line 11; also see line 10). We learn that the accident is linked to the Princess being “pursued by paparazzi” and that one of Europe’s strictest privacy laws – in France – did not prevent this incident (see line 3). This line appears in an article titled “Why paparazzi are beyond the law” (Bodkin, 1997), which argues that the law only allows celebrities to sue the publications rather than the photographers for a fine and that all involved parties benefit financially (including the celebrities, if their court cases are successful). Another argument against introducing a privacy law is given in line 8, cautioning against the risk of introducing a law in the “emotion-charged aftermath” of the accident, which the writer argues would be “deeply-flawed” (from the extended context of the line in a letter to the editor: Connew, 1997).

Following the accident, another dimension of privacy concerns is voiced in relation to the sons of Princess Diana, the *princes* (a significant 1997 collocates in Figure 6-11). As children at the time of their mother’s accident, they receive particular media attention. Unlike the concordance lines focusing on the Princess, Concordance 6-5 shows a much more uniform picture of broad agreement about protecting the privacy of the princes. This impression is raised through modals: “must be protected” (line 3) or “respected” (line 5) and “would not [...] let the privacy be violated” (line 4). In addition, the lines emphasise that the privacy they are granted

will be complete (“devise a blanket of total privacy”). Various news outlets are quoted giving their “firm and absolute assurance” (line 6) to protect the privacy or that they have no intention to publish photos of the princes (line 8). Although there is no mention of privacy law in this context, line 7 refers to a press “code” and the possibility to extend it to protect children throughout their “full-time education” (also see line 5). Overall, Concordances 6-4 and 6-5 demonstrate that at this point in the discourse, on the one hand, the introduction of a new privacy law protecting celebrities is viewed as controversial. On the other hand, there is widespread agreement on the need for protecting the privacy of (celebrities’) children.

- [1] Prince, of Wales are on friendly terms; they could together --- devise a blanket of total **privacy** for the *Princes*, and even --- extend it with reasonable efficiency to foreign publications. If it
- [2] to walk and to do the things teenage boys would --- do in this situation in **privacy**." The *Princes* will travel to --- London today with their father. By Nicholas Watt German Language
- [3] allowed of members of the family while they were giving --- readings or speeches, but the **privacy** of the two *Princes* and --- the rest of the family must be protected while seated
- [4] of the modern age", and vowed that the family would --- not let the two *princes*' **privacy** be violated in the same --- way. Moves to increase the protection of children were
- [5] 22,23 children , Damian Whitworth says Lord Wakeham addresses the London --- Press Club yesterday. The *princes*' **privacy** must be respected throughout their --- education, he said University of Oxford \$^\\ University of W Oxford
- [6] Sunday Mirror, gave readers "a firm and absolute assurance --- that we will respect the **privacy** of the young *Princes*". The --- paper would work closely with the commission to ensure that
- [7] and Prince Harry until they are 16. The code may --- be amended to protect the **privacy** of the *Princes* and other --- children until they have finished full-time education . The committee
- [8] their conduct : "The Sun has no intention of --- carrying photographs which invade the **privacy** of *Princes* William and Harry." --- Among the tabloids , the most prominent exponent of a tighter

Concordance 6-5: All 8 examples of *privacy* co-occurring with *princes* in the 1997 subcorpus, shown with an extended context (total span of 15)

6.2.2.3 *Privacy collocates in 2008*

The 1997 focus on introducing laws for privacy protection is absent from the 2008 collocates. On the contrary, they appear to show some of the outcomes of privacy regulation, suggesting that by this time privacy has been institutionalised: many 2008 *privacy* collocates shown in form part of the “small print” section in various special reader offers. Example (11) illustrates

the formulaic mention of the privacy policy that repeatedly occurs in these ads. The asterisks highlight the collocates from Figures 6-11 and 6-12.

(11) Tick if you don't want to receive these from us [] or *carefully* *selected*
companies [] (see our **privacy** *policy* at *www.nidp.com*).
("The Times," 2008)⁷

Finally, in a similar manner to *privacy* in the 1986 housing ads, the 2008 collocate *glass* depicts privacy as a valuable asset and is here used to promote cars. Arguably, a car is the next most personal space, making privacy a desirable feature. In 2008, car adverts from a number of brands (e.g. Toyota, Nissan, Land Rover, Ford) mention (*rear*) *privacy glass* as standard fitting of the models. The online interface yields no results for the compound in 1986; likely, the technology was only developed afterwards.

6.2.3 Long term comparisons for *CCTV*

By contrast to the general concepts of *surveillance* and *privacy*, the node *CCTV* is more dependent on technological development. This development is reflected in the frequency profile of the term (see Section 6.3): *CCTV* is infrequent in the 1980s and its usage only increases in the late 1990s. The following examples of collocates from the years 1986, 1997 and 2008 illustrate some of this development in terms of a shift in the discourse. For *CCTV*, the long-term comparisons only produce two result plots (Figures 6-14 and 6-15), as there are no significant results for the comparison of 1986 vs. 1997 in *CCTV* collocates.

⁷ This has been typed from the image of the online interface.

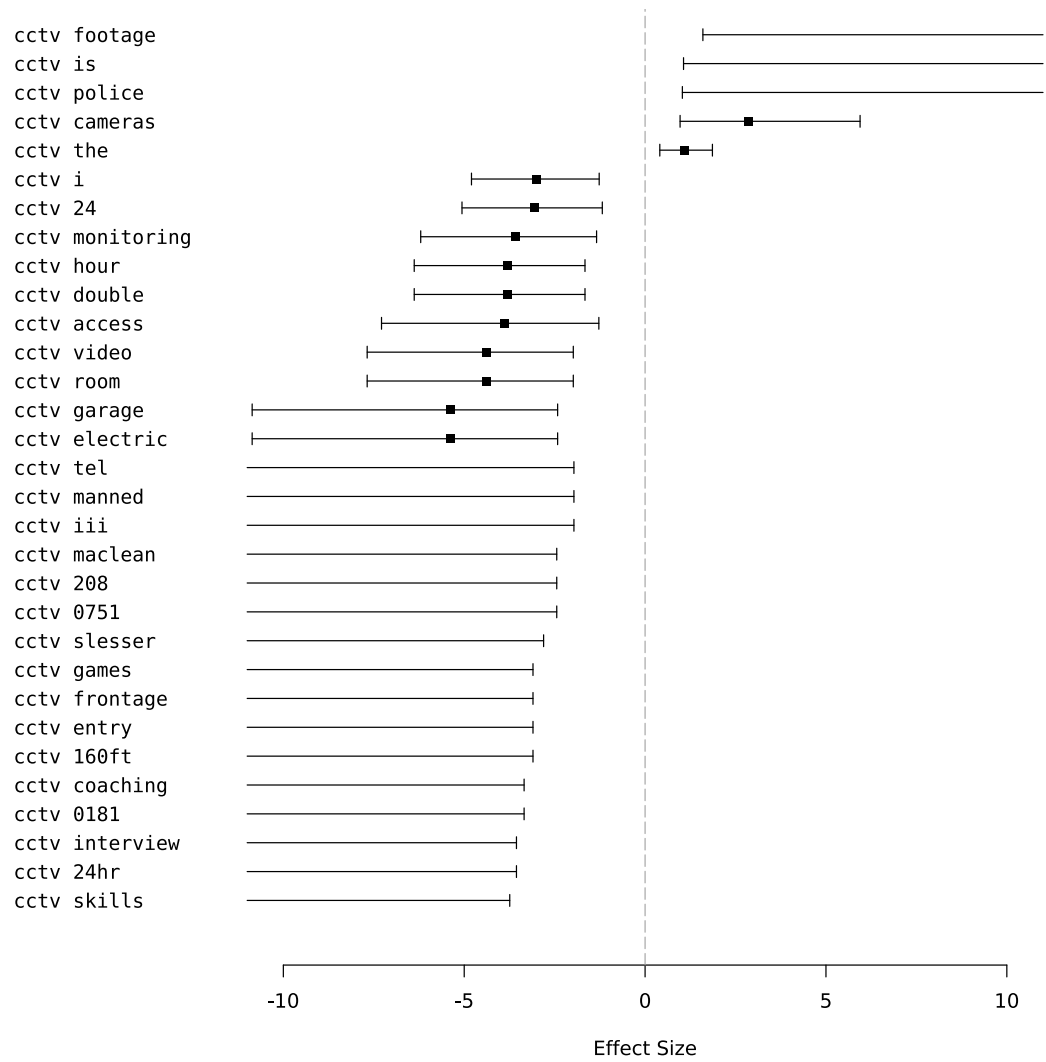


Figure 6-14: Co-occurrence comparison of *CCTV* in the whole year 1997 (left) vs. 2008 (right)

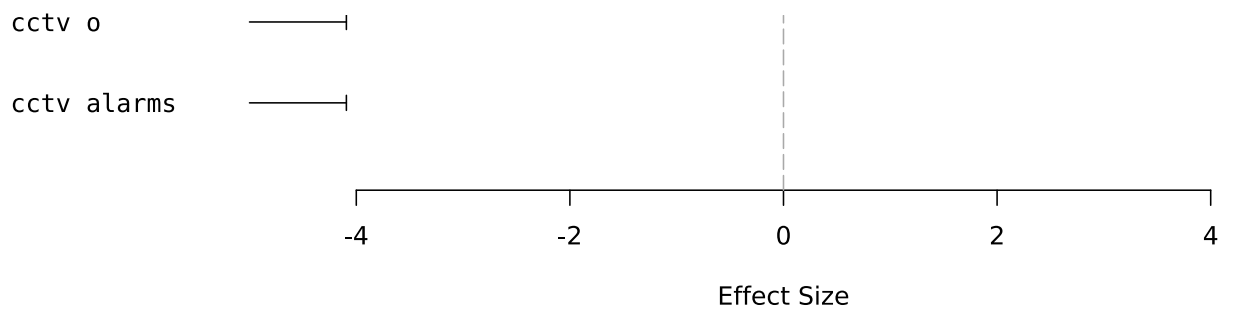


Figure 6-15: Co-occurrence comparison of *CCTV* in the whole year 1986 (left) vs. 2008 (right)

6.2.3.1 CCTV collocates in 1986

In the year 1986, *CCTV* only occurs five times. This lack of occurrences in 1986 is likely the reason that no significant results are found for the first comparison of 1986 vs. 1997. When comparing 1986 vs. 2008 for *CCTV*, two significant results are found (see Figure 6-15). Given the low frequencies of *CCTV* in 1986, the results are based on low co-occurrence frequencies. Both collocates originate from classified ads from security companies: one advertises a job vacancy (Figure 6-16; “First Inertial Systems Limited,” 1986) and the other one promotes security services and equipment (Figure 6-17; “Multiple Display Advertising Items,” 1986). The OCR scan has recognised the bullet points as the letter *o* that is visible in Figure 6-15). These two ads represent technical mentions of *CCTV* that assume that readers are familiar with this new technology. This knowledge cannot necessarily be assumed for non-technical readers, as an example from the TV programme in 1997 still spells out the term (see the next subsection).



Figure 6-16: 1986 job ad containing *CCTV*



Figure 6-17: Ad for security services containing *CCTV*

Apparently, the innovative use of *CCTV* as a new technology is first mentioned in advertising: four out of the five 1986 mentions appear in adverts, while the one instance in news refers to the Chinese state TV channel “China Central Television” (“Chinese TV turns back the pages,” 1986). Accordingly, in 1986, *CCTV* as a surveillance technology does not occur in news articles in 1986, but only in technical specifications of classified ads. At this point, there is no debate

about the benefits or risks of this technology in the News section. Its implications for society are not discussed. Section 6.4 focuses on the relevance of these distributions across the newspaper sections for all three nodes.

6.2.3.2 CCTV collocates in 1997

When comparing 1997 to 2008, a number of significantly different collocates are found (Figure 6-14). In a similar way to the 1986 patterns, the results on the side of 1997 are dominated by advertising, in particular, the interview skills service ad shown in Figure 6-18 and the housing ad in Figure 6-19. Many of the collocates from the 1997 side of Figure 6-14 feature in the direct proximity of *CCTV* in these ads: *interview*, *skills*, *coaching* as well as the proper nouns *Slessor*, *Maclean* and part of the phone number in Figure 6-18 (“Applying for Jobs,” 1997) and *games*, *room*, *double*, *garage*, *24hr*, *video*, *entry*, *160ft*, *frontage*, *electric* in Figure 6-19 (“Hadley Wood,” 1997). In both cases, CCTV is used as a valuable asset, an innovative technology in the *interview skills coaching* workshop and one of the outstanding features of the *luxury house* to provide *24hr* security. Since space is limited in classified ads, the inclusion of *CCTV* implies that it is considered an important feature for promotion in both 1986 and 1997.

**APPLYING FOR JOBS ? Through
Career change or internal pro-
motion. Enhance your chances
of success by interview skills
coaching using CCTV Slessor
Maclean 0181 208 0751**

HADLEY WOOD
Luxury house set in approx
1 acre, 8 bedrooms, 5 en
suite, 6 receptions, gym,
utility, games room, double
garage. 24hr CCTV, video
entry. 160ft frontage,
electric gates. £1.3 million.
0171 247 6361.

Figure 6-18: Ad for interview skills coaching containing CCTV

Figure 6-19: Housing ad containing CCTV

Many of the ad-related collocates are unique to 1997, as those same ads do not feature anymore in the 2008 subcorpus. An example of a 1997 collocate that is not unique is *monitoring*. As illustrated in Figure 6-20, the description of the episode for BBC series *Crime Beat* shows a snapshot from the adoption phase of CCTV (“BBC 1,” 1997). At this point, the name of the technology still needs to be spelled out before it can be abbreviated in the second sentence. The episode focuses on the “success of CCTV monitoring”, and although *Big Brother* is mentioned, it is used in connection with a positive statement (“helping to eliminate computer theft”). The frequencies that these results are based on are still low, but they are higher than in 1986 (also see Section 6.4). This is evidence that the discourse around *CCTV* begins to mature as discussions of its use in fighting crime start to appear, even if this takes place in the TV programme rather than in feature articles.

8.00 Crime Beat Martyn Lewis explores the growing popularity of closed-circuit television for reducing crime. He visits Accrington in Lancashire, where the success of CCTV monitoring has greatly reduced crime in public places, and learns how the “Big Brother” approach is helping to eliminate computer theft in business parks. Plus, the latest anti-theft gadget is unveiled — the smoke barrier (T) (7609)

Figure 6-20: TV programme description with the co-occurrence pair (*CCTV*, *monitoring*)

6.2.3.3 *CCTV* collocates in 2008

This section focuses on *CCTV* collocates in the final year of the corpus. My examples show that by this point, the *CCTV* technology has existed long enough for it to form part of the public discourse. Thus, by 2008, the co-occurrence patterns suggest that *CCTV* has become mainstream, now clearly including a large number of examples from news articles. The following examples show that *CCTV* is partly taken for granted in certain text types such as

articles reporting on crime. Unique 2008 collocates of *CCTV* compared to 1997 are *footage*, *police* and *is*. The first two collocates in particular tend to feature in reports of (criminal) instances as in Examples (12) and (13).

(12) *Police* later released **CCTV** *footage* of Omar, who is 6ft 2in tall, dressed head to toe in a burka. He had shaved his arms in an attempt to look more feminine.

(Bird, 2008)

(13) *Police* yesterday released **CCTV** *footage* of the incident, which happened on May 23.

(Mostrous, 2008c)

The collocate *is* appears in various constellations with *CCTV*. Some of these instances do not directly describe CCTV or its features. In Example (12), *is* forms part of the relative clause “who is 6ft 2in tall” and therefore contributes to the function of CCTV footage to provide evidence for an ongoing crime case, in this case by specifying the personal appearance of the suspect. In other cases, a form of BE is used in more general discussions about whether or not CCTV is *protecting us* (see the heading shown in Example 14). Similarly, occurrences of *CCTV* with *is* that discuss its efficiency and importance are shown in Example (15), which contains extracts from an article on CCTV and other surveillance measures. The first sentence in (15) also includes a second 2008 *CCTV* collocate from Figure 6-14, forming the compound *CCTV camera*. These examples indicate that the discourse on CCTV has evolved to the extent that it reflects on the technology, its value, personal reactions and its social implications. Example (15) explicitly makes connections with notions of unease (“slightly chilling”) and with surveillance imagery (“Big Brother”). At the same time, the latter part of the extract argues that “CCTV is an invaluable resource”.

- (14) Safe and secure? **CCTV** *is* not protecting us, say police chiefs
(“Index,” 2008)
- (15) The modern **CCTV** *camera* *is* slightly chilling in its efficiency. It can pan from side to side and zoom in to a far higher resolution than you might expect. [...] Big Brother may be watching us, but he doesn’t seem very interested. [...] **CCTV** *is* an invaluable resource for surveillance, but cameras are not (yet) in every corner of the land.
(Rifkind, 2008)
- (16) The ‘Town Hall Stasi’ have been pilloried by politicians for tracking our every move but, as Hugo Rifkind discovers, the eyes behind the surveillance cameras can be a force for good – tackling everything from noisy neighbours to dog mess.
(Rifkind, 2008)
- (17) You may find all of this rather chilling, in a country such as ours. You may wonder if an Englishman’s home will ever be his castle again. Me, I just think of that weary couple, not wanting any more trouble, just wanting some sleep. And, you know, I’m not that fussed.
(Rifkind, 2008)



Figure 6-21: Pages 2–3[S] in *times2* supplementary material of *The Times*, 21 November 2008

The article from which Example (15) is taken argues that the benefits ultimately outweigh any concerns. This is indicated in its lead paragraph, reproduced in (16), which gives examples for surveillance being “a force for good”. And although the main case study is unrelated to CCTV – it reports on a council installing “a bug”, in response to a couple’s complaints about their next-door-neighbour – cameras feature as prominent symbols of “Big Brother” on the page

spread (Figure 6-21; Rifkind, 2008). Despite the somewhat intimidating look of the CCTV cameras, the use of *only wants to help* in the heading connects with the concept of surveillance as “a force for good” in (16). This clash of the negative and positive connotations of surveillance appears to be deliberately provocative. Similarly, the final paragraph, shown in (17), picks up on the question whether the surveillance developments are *chilling*, but ends on the note that there is no reason for concern (“I’m not that fussed”), arguing that the measures do more good than harm.

6.2.4 The development of reflective public discourses on surveillance

This section has addressed RQ 3-1, by analysing long-term changes in co-occurrence patterns of the surveillance discourse. In comparison to 2008, the earlier years have tended to reveal more patterning in advertising than in news articles. Yet, a reflection of social values has to some extent transpired from the co-occurrence results of all subcorpora under study. Co-occurrence pairs like (*privacy, law*) in the 1997 subcorpus have indicated a deep engagement with questions of privacy, partly as a result of the events surrounding the tragic accident of Princess Diana. The 2008 co-occurrence patterns of *CCTV* reflect more complex discussions of surveillance than seemed to occur in the results from the earlier subcorpora. At the same time, the compound *surveillance society* reflects the development of a genuine debate over surveillance measures in 2008.

6.3 Co-occurrence patterns at monthly peaks

This section moves from the long-term focus of 6.2 to more localised patterns. For a bottom-up selection of the subcorpora, this section begins with a frequency overview. Figures 6-22 to 6-24 show the frequency developments of *surveillance*, *privacy* and *CCTV* across the

TDA1986–2008. In each plot, the top five peaks in relative frequency are labelled with their year and frequency per million words.

According to these figures, both *surveillance* and *privacy* occur throughout the period and their relative frequencies appear to rise gradually, despite some variation in the middle of the period, with the highest relative frequency peak of *privacy* occurring in January 1993. The frequency profile of *CCTV* differs in that the word hardly occurs in the first half of the corpus period. Its relative frequency then increases with sudden peaks in the late 1990s and 2000s. This trend may reflect the UK government’s introduction of open street CCTV campaigns for crime prevention in the 1990s that was mentioned in Chapters 2 and 5 (also see Norris et al., 2004).

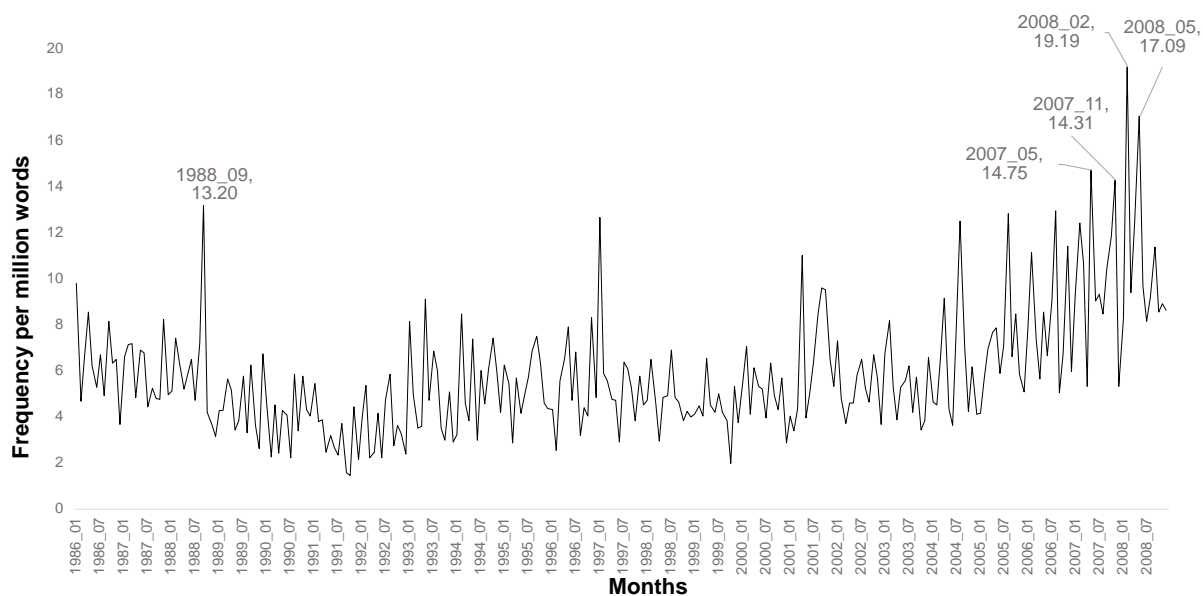


Figure 6-22: Relative frequency of *surveillance* per month, TDA1986–2008

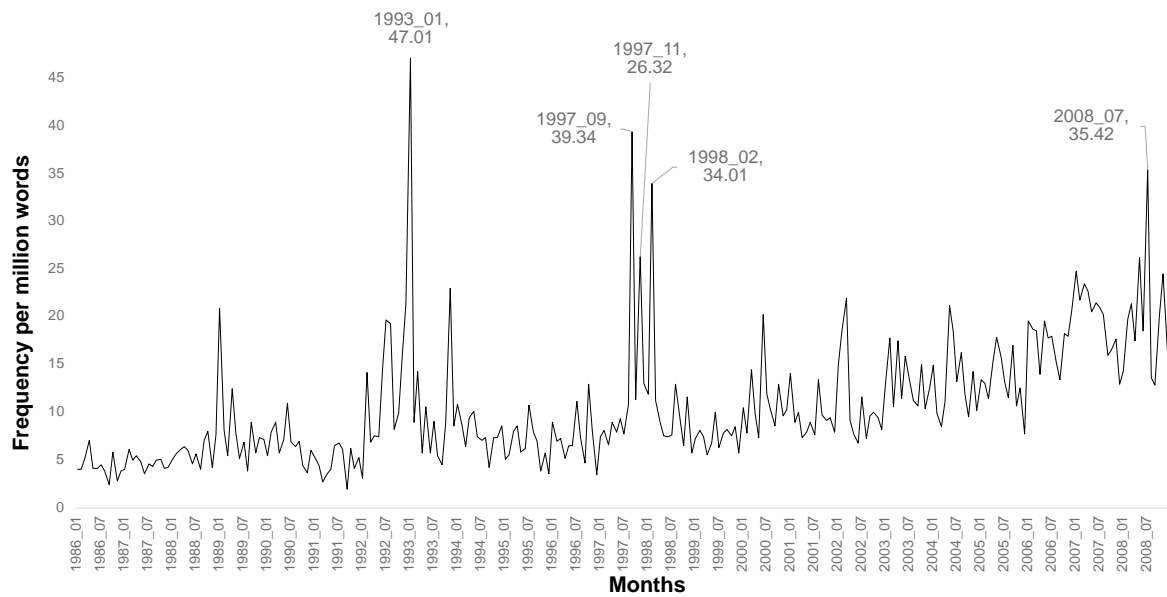


Figure 6-23: Relative frequency of *privacy* per month, TDA1986–2008

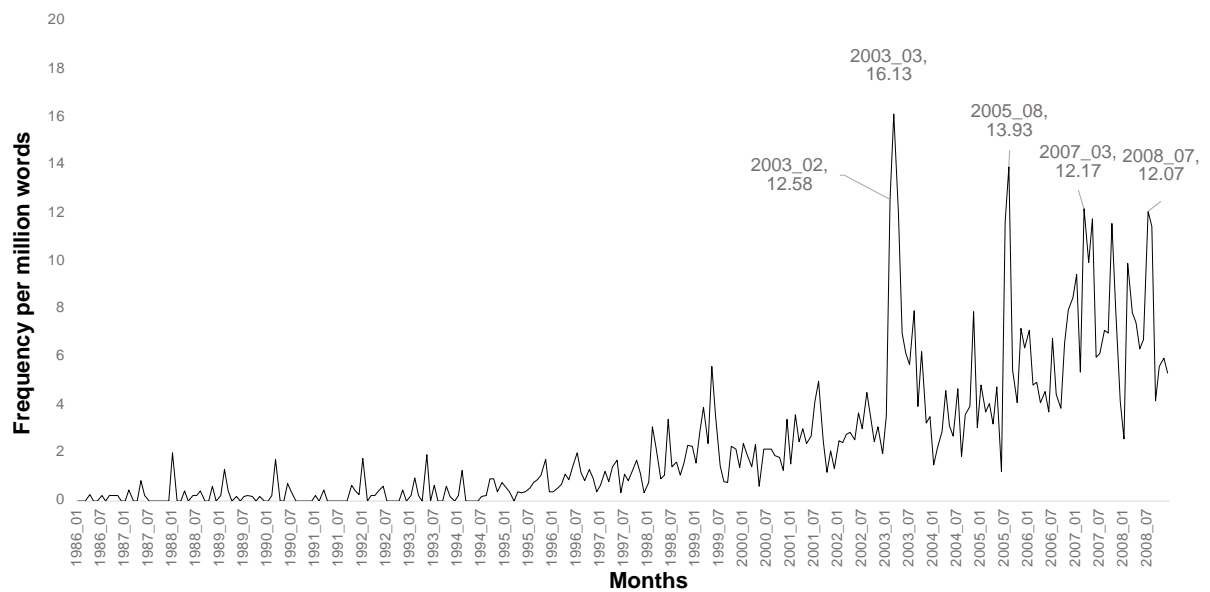


Figure 6-24: Relative frequency of *CCTV* per month, TDA1986–2008

These peaks indicate when the respective node is comparatively mentioned the most. Increases in relative frequency of particular words can point to overall shifts in the discourse. For example, Zinn and McDonald (2017) begin their study of risk discourse in *The New York Times* by tracing the changes in relative (and absolute) frequency of a group of risk words. They work

with annual subcorpora, as they are more concerned with overall trends than with peaks triggered by particular events, which are the focus of this section.

The type of corpus also plays a role in choosing the method. In the TDA1986–2008, surveillance discourses only represent one among many other concerns. As a result, the word *surveillance* is a comparatively infrequent content word (as opposed to *government*, for instance).⁸ Accordingly, I chose peaks in the relative frequency in order to identify salient patterns in this general corpus. When working with more specialised corpora, it may be useful to use a different definition of what constitutes a peak. For example, Gabrielatos et al. (2012, pp. 164–165), who work with a large specialised corpus consisting of newspaper articles that contain a query term related to Islam and Muslims, identify peaks by looking for the biggest frequency rise from a trough.

For the co-occurrence comparisons with each peak, I take the first full corpus year, 1986, as the reference corpus. Only for *CCTV* I use the year 1999 as the baseline instead, because the word was infrequent prior to that (see Figure 6-24). As the aim of this section is to investigate local patterns at peaks in the frequency profiles of the nodes, I only show and discuss the co-occurrence patterns of these peak months, i.e. the right side of the plots, and omit the results for the full year reference corpora.

6.3.1 Peak comparisons for *surveillance*

In comparing the collocates of *surveillance* in their peak months with a baseline year, the assumption is that the peaks exist because of newsworthy events related to surveillance discourse that prompted more coverage. The results for the five peaks are given in Figure 6-25.

⁸ The average monthly relative frequency of *surveillance* from 1986_01 to 2008_12 is 5.86 per million; for *government*, it is 363.85 per million.

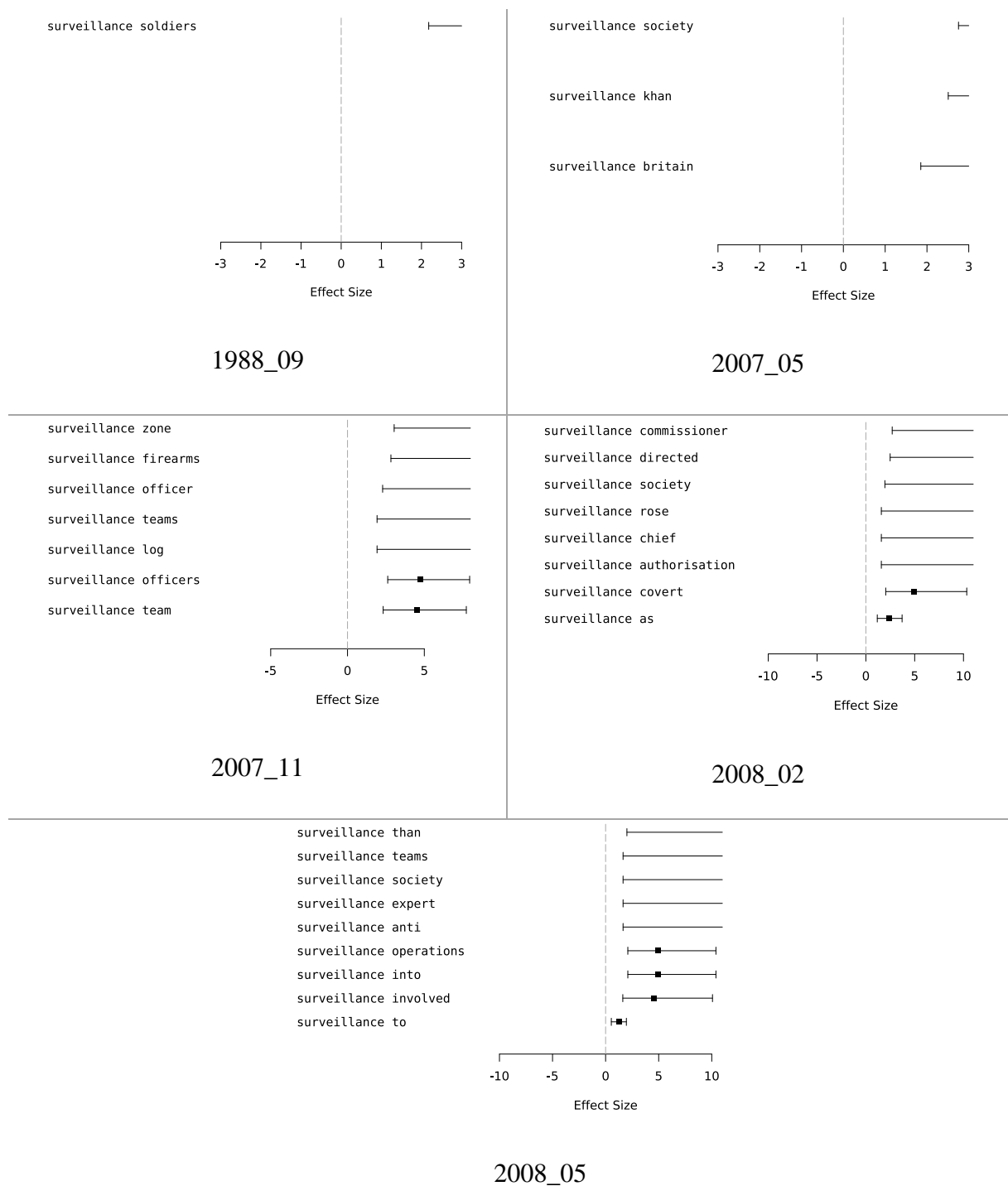


Figure 6-25: Co-occurrence comparison of *surveillance* in each of the top 5 peak months (vs. 1986), ordered chronologically

Chronologically, the first peak appears not long after the 1986 baseline year, in 1988_09. The only result for this peak (see the top left panel of Figure 6-25) is the collocate *soldiers*, which

is unique in this peak month. As this collocate does not co-occur with *surveillance* at all in 1986, the comparison picks up this pair at a relatively low co-occurrence frequency: all five instances are shown in Concordance 6-6. The lines mostly refer to *soldiers* as having separate duties from “operatives” (line 3) and “teams” (line 5) carrying out surveillance, although in line 2 the soldiers themselves are “on surveillance at the border” in response to “terrorists”. A search in the TDA interface indicates that lines 2–5 originate from the coverage of the “Gibraltar Inquest” to investigate the circumstances of a military operation by the British Special Air Service (SAS, see line 5) during which members of the IRA (“Irish Republican Army”) were killed. Line 1 is also related to the Northern Ireland conflict, but refers to conflict conditions in Northern Ireland itself.

```
[1] to carry --- out its key tasks of surveillance , and transporting weapons and *soldiers*
[2] - One involved armed *soldiers* on surveillance at the border as the --- terrorists
[3] --- other teams of *soldiers* and surveillance op era tives on the --- ground - "a
[4] itish intelligence involved in the surveillance operation, and then by *Soldiers* --- A t
[5] ld then --- have been picked up by surveillance teams and the SAS *soldiers* --- would
```

Concordance 6-6: All 5 instances of (*surveillance*, *soldiers*) in 1988_09

The remaining peaks are more clustered together at the other end of the corpus, all occurring in 2007 and 2008. A recurrent collocate is *society*. It appears in three plots in Figure 6-25, to which I return at the end of the subsection. One example of a local co-occurrence pair is (*surveillance*, *khan*) from the comparison results of the 2007_05 peak (top right panel in Figure 6-25). Concordance 6-7 shows that *Khan*, in this context, is the name of a terror suspect of the “7/7 bombings” on London underground trains in 2005. Even this short concordance creates the impression that the articles question the surveillance strategy of the secret service (MI5, but misrecognised as *MIS* by the OCR output in line 4), as apparent from “what happened with the surveillance pictures” and the question “was that mistake avoidable”. This depiction of the situation is explained by the wider context, as the articles report on an ongoing investigation

within the secret service, shown in Example (18), where “ISC” stands for “Intelligence and Security Committee”. These instances probe whether surveillance measures have not been enacted properly or sufficiently, and, therefore, might have missed an opportunity to avoid the attacks.

[1] made --- the *right judgment* in switching **surveillance** away from *Khan* to other ---
 [2] attack. In --- the case of *Khan*, this **surveillance** extended back to 2003 and --- more
 [3] t Khyam four --- times when he was under **surveillance** in early 2004 were *Khan* --- and
 [4] ecurity --- service. Yet MIS *dropped* the **surveillance** on both *Khan* and Tanweer, ---
 [5] on the car --- and addresses and set up **surveillance** on *Khan* and * Tanweer, which ---
 [6] *disarray* --- over *what happened* with the **surveillance** pictures of *Khan* and Tanweer. ---
 [7] he *wrong* --- decision to take *Khan* off **surveillance**. *Was that mistake avoidable?*

Concordance 6-7: All 7 instances of (*surveillance*, *Khan*) in 2007_05

(18) The ISC’s other main focus will be whether MI5 made the right judgment in switching surveillance away from Khan to other suspects.

(O’Neill & Coates, 2007)

A different type of surveillance is highlighted by the collocate *zone* which is revealed in the comparison of 1986 vs. *surveillance* peak month 2007_11 (middle left panel in Figure 6-25). This co-occurrence arises from the concept of “surveillance zones” in the fight of livestock diseases. Example (19) illustrates the use of this term in reference to a large-scale outbreak of avian flu in November 2007. This instance clearly portrays a different context of surveillance, where privacy is irrelevant, and instead the economic loss in the livestock “market” takes centre stage.

(19) East Anglia produces about a third of Britain’s turkeys and there are fears that if the virus takes hold the £400 million market for Christmas birds will be lost. There are three million birds alone in the 10km (six-mile) **surveillance** *zone* around Redgrave Park Farm.

(Elliott & Duncan, 2007)

The third and final example of collocates that are specific to one peak involves several *surveillance* collocates from the 2008_02 peak period (see the middle right panel in Figure 6-25). Example (20) shows that they co-occur in the text. The example reports on an investigation into the “bugging” of a conversation between MP Sadiq Khan⁹ and an inmate at a prison. The article is titled “Illegal bug uncovered in second UK prison” and suggests that covert recording is “widespread”, framing this surveillant measure as a serious problem. This example illustrates that surveillance has by now been officially recognised as a social issue that merits the official appointment of a *Chief Surveillance Commissioner*. So, this title is another example of an institutional name that contributes to the cultural keyword status of *surveillance* (see Section 6.2.1). Because the commissioner is mentioned in this and further articles related to the prison “bugging” story, the collocates highlighted in (20) are significantly more frequent in 2008_02 than in 1986.

(20) Mr Straw, the Justice Secretary, told MPs yesterday that Sir Christopher *Rose*, the *Chief* **Surveillance** *Commissioner*, is to head an inquiry into the bugging of a conversation involving Mr Khan and Babar Ahmad when the MP visited him in Woodhill jail.

(Ford & O'Neill, 2008)

The shared collocate *society* links back to the rise of the bigram *surveillance society* that I discussed in Section 6.2.1, where I found that this term occurs most frequently in the last two years of the TDA1986–2008. From a methodological point of view, it is reassuring that the separate long-term and peak comparisons reveal these similar results. Furthermore,

⁹ The *Khan* mentioned here is not the same person as the one referred to in Concordance 6-7 (and *Khan* is not a significant collocate for 2008_02). This demonstrates that it is useful to restrict concordance lines to both particular temporal subcorpora and a selection of collocates, as the same surface form is more likely to refer to the same context if the temporal period is limited.

concordances show that other collocates from these peaks interact with the concept of the surveillance society. Example (21) illustrates this connection for the surveillance collocate *operations* from the 2008_05 peak. Although the bigram *surveillance society* does not appear in the near proximity of this co-occurrence pair, a similarly critical connection is made to the “hallmarks of dictatorships long ago”. Yet, in (22), an example from the 2007_05 peak, *Britain* co-occurs with *surveillance society*, emphasising that the label refers to the current circumstances.

- (21) Thousands of council workers across Britain today routinely authorise and engage in **surveillance** *operations* that were the hallmarks of dictatorships a generation ago.
(“The Lives of Others,” 2008)

- (22) His organisation, which launched a report last November branding *Britain* as a **surveillance** *society*, with individuals likely to be filmed by 300 cameras every day [...]
(Reid, 2007)

As a final example, consider (23), which originates from a surveillance-themed double-spread published on 31 May 2008 (see Figure 6-26). In a similar manner to Example (22), this article relates to *surveillance operations* carried out by local councils. This debate appears to be triggered by an investigation into how the law on surveillance operations is implemented locally.

- (23) Amid increasing concern in Parliament that the UK is slowly becoming a **surveillance** *society*, the committee has looked at the operation of the Regulation of Investigatory Powers Act (Ripa), which some MPs say is being misused to focus on petty crime rather than serious offending.
(Mostrous, 2008b)

of council officials into spies

But the surveillance must be "necessary and proportionate" and it is those two words the critics are fixated on. The Home Office says it is really necessary to use such powers to deal with drug dealing or terrorism.

Charities groups think councils are being heavy-handed, and the Act intended to provide protection for the right to privacy is now being used to strip it away, satisfying the right to a private life.

→ The Home Office says as

The spread is framed as evidence for the country increasingly turning into a *surveillance society*. The phrase is used not only in the text of Example (23), but also in the header of both pages (Figure 6-26). The overall title and the headings of two shorter articles all appear to question whether the surveillance measures are disproportionate in relation to the problems they are meant to resolve (such as “dog fouling”) or the vulnerability of the groups being watched (e.g. “children”). To some extent, these instances therefore resemble the sentiments expressed in the context of *covert surveillance* in Section 6.2.1.

300

6.3.2 Peak comparisons for *privacy*

Figure 6-27 shows the relative frequency of *privacy* in the TDA1986–2008. Like *surveillance*, *privacy* is never absent from the monthly periods. There is considerable overlap between the co-occurrence results for the peak months and the patterns that emerged from the long-term comparisons. This is not surprising as three out of the five top *privacy* peaks occur in the years that were compared in Section 6.2.2 – 1997 and 2008.

Several of the peaks link to the theme of law. Concordances show that these law-related peaks fall into two groups: the first peak, both in terms of its relative frequency ranking and chronologically is 1993_01 and a second group of the peaks from 1997 to 1998. Both groups deal with discussions about privacy infringements by the press. Concordance 6-8 displays all instances of *privacy* co-occurring with *civil* in the 1993_01 subcorpus. It indicates that at this point, the discourse of *privacy* is concerned with the potential introduction of a “civil law of privacy” (see lines 1, 2, 10, 11), saliently highlighted by the recurring adjective *new* (lines 1, 2, 5, 6, 8, 10). Many lines contain modals to discuss the potential existence of the law: e.g. *possible*, *could* (line 1), *would* (lines 4, 12), *will* (line 10) and the conjunction *if*, which itself features as a significant 1993_01 collocate of *privacy* in Figure 6-27. The deontic modal verb *should* in line 14 expresses the view that civil law is indeed the right branch of law to handle privacy intrusion.

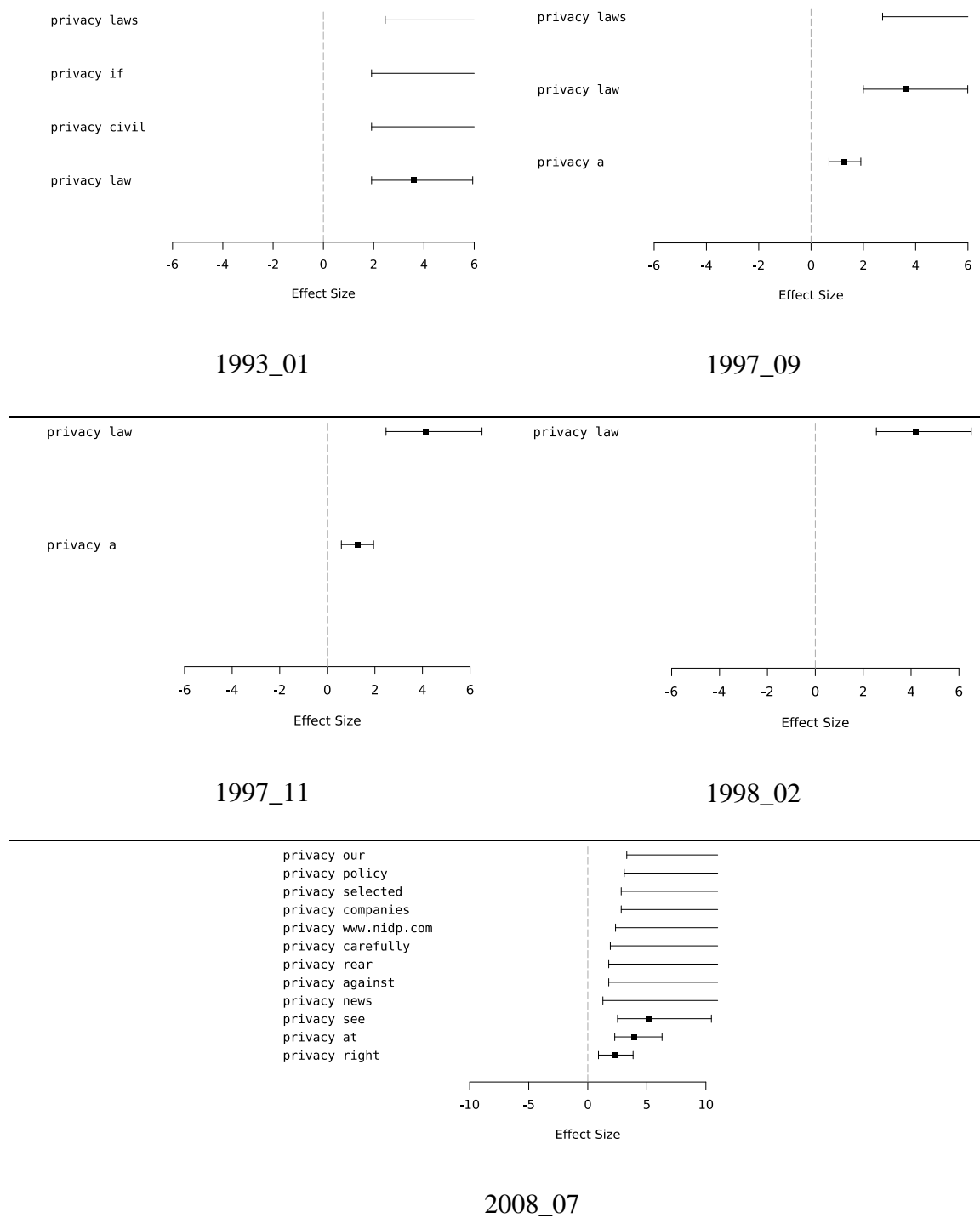


Figure 6-27: Co-occurrence comparison of *privacy* in each of the top 5 peak months (vs. 1986), ordered chronologically

[1] possible new *civil* law of **privacy** if one could be clearly
[2] a new *civil* law of **privacy** and leave it up to
[3] *civil* action for invasion of **privacy** just as they can now
[4] *civil* law of breach of **privacy** that would allow individuals to
[5] *civil* remedy for people whose **privacy** was infringed. The new criminal
[6] backs laws to protect **privacy** NEW criminal and *civil* laws
[7] *civil* laws to protect the **privacy** of individuals were heralded
[8] 8 million). Brooke promises new **privacy** laws New criminal and *civil*
[9] *civil* laws to protect individuals' **privacy** were heralded by the government
[10] any new *civil* law of **privacy**, will face tough opposition from
[11] oppose a *civil* law of **privacy** if it was considered as
[12] new *civil* offences, to protect **privacy**. Mr Brooke said legislation would
[13] a law of infringement of **privacy** as a *civil* right of
[14] Generally, matters of intrusion into **privacy** should be left to *civil*
[15] *civil* wrong) of infringement of **privacy**. They said that if Mr

Concordance 6-8: All 15 instances of *privacy* co-occurring with *civil* in 1993_01

Further evidence for the negotiation of the discourse being in progress is given by lines with contrasting views. On the one end of the spectrum, we find implied support for introducing a new law (line 6) or the impression that the success of the law is taken for granted (“promises”: line 8; “heralded”: line 9). On the opposite end there is “tough opposition” (line 10; also see “oppose”: line 11). Two lines (8 and 12) refer to the then National Heritage Minister, (*Mr*) *Brooke*. These references link the discussion of the potential new civil law to statements by Brooke in response to a report published in January by Sir David Calcutt, the “Review of Press Self-Regulation” (see Bingham, 2007 for an overview of this “crisis” in the self-regulation of the British press). The January peak therefore indicates an ongoing debate about privacy that feeds into political decisions.

As for 1997 as a whole (see Section 6.2.2), the collocates *law* and *laws* are found to be significantly more frequent collocates for the peak comparison of 1997_09, just following the death of Princess Diana on 31 August 1997. They still appear in the results of both 1997_11 and 1998_02 in the new year. Discussions of introducing a *law* do not feature any more in the results for the 2008_07 peak. Instead, we find a situation where people have a *right to/of privacy*, apparently triggered by various distinct cases relevant to the right to privacy (see Concordance 6-9). Accordingly, the “right to privacy” by 2008 is “fundamental” (line 17), it is “derived” from (line 19), “bestowed” (line 4) or “provided” by law (line 20). People “have”

(line 15) a right to privacy, but this right can be LOST (line 1), BREACHED (lines 7, 14), WAIVED (line 8) and “ignored” (line 13). Although the right to privacy seems more institutionalised than in the earlier years, negotiation continues, for example in relation to how far the right “extends” (line 6).

```
[ 1] f YouTube --- users lose *right* to web privacy Holden Frith Google must divulge --- the
[ 2] -- but heavily camouflaged." *Right* to privacy 'should not extend to violence' ---
[ 3] obstructed --- by the human *right* to privacy, the News of the World's --- counsel told
[ 4] hts, --- which bestowed the *right* to privacy, had been drawn up in --- the immediate
[ 5] o it, Mr --- Warby said. The *right* to privacy, which enshrines respect for human ---
[ 6] ill --- consider whether the *right* to privacy extends to sadomasochism and paid---- for
[ 7] ed --- breaching Mr Mosley's *right* to privacy, however. Uprising at M&S ---
[ 8] dren in --- care waive their *right* to privacy if they wish to speak --- out. For
[ 9] nsider --- whether the human *right* to privacy should include sado-masochistic and ---
[10] er, James --- Price, QC. The *right* of privacy is a new feature of --- English law,
[11] balanced --- with a child's *right* to privacy". But TIMÂs TIMES the --- two should not
[12] balanced --- with a child's *right* to privacy. These are sensitive issues and --- it is
[13] d --- recklessly ignored his *right* to privacy and was only interested in --- the
[14] oors: does --- it breach the *right* to privacy or is it the media's --- right under
[15] that they --- have a *right* to invade privacy minute by minute. "We've Come --- across
[16] the eyes --- ofthe law, the *right* to privacy at work is more far---- reaching than
[17] octrine --- of a fundamental *right* of privacy, that could not be found --- in the US
[18] ution: "We --- deal with the *right* of privacy older than the Bill of --- Rights â "
[19] erbally. --- He derived this *right* of privacy from the US Bill of --- Rights, where it
[20] ourt, and --- provided a new *right* of privacy parallel, though subordinate, to the ---
[21] nd popular --- support for a *right* to privacy victims ofthe Holocaust. It may --- have
[22] k effect in --- 2000 and the *right* to privacy under Article 8 was enshrined --- in
[23] Convention --- rights of the *right* to privacy under article 8 and light --- to freedom
```

Concordance 6-9: All 23 instances of (*privacy*, *right*) co-occurring in 2008_07

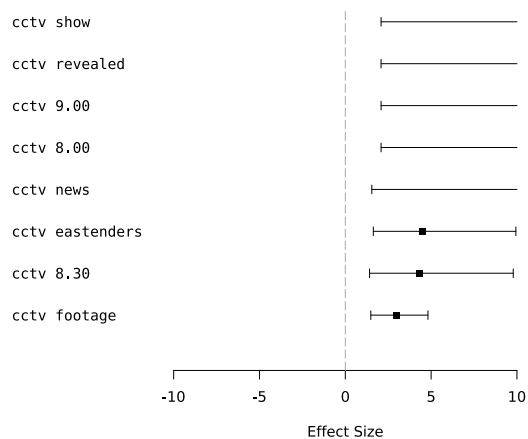
(24) But there was support for tightening the laws of trespass to stop snooping and bugging, as well as for a possible new *civil* law of **privacy** if one could be clearly defined.
(Gibb, 1993)

Co-occurrence patterns of *privacy* that have emerged from the long-term comparisons in 6.2.2 and for the peaks mostly relate to discussions of laws that protect privacy and infringement thereof. The types of privacy violations that have featured saliently in this data are mostly related to disclosures of an individual’s personal information by the media. These cases are not usually framed in terms of surveillance, but are discussed in the light of balancing individual privacy and freedom of press. Nevertheless, some of the *privacy* peaks appear to interact with surveillance discourses. Example (24), the extended context of line 1 in Concordance 6-8, illustrates such a case. The new civil law of privacy is mentioned together with an initiative to

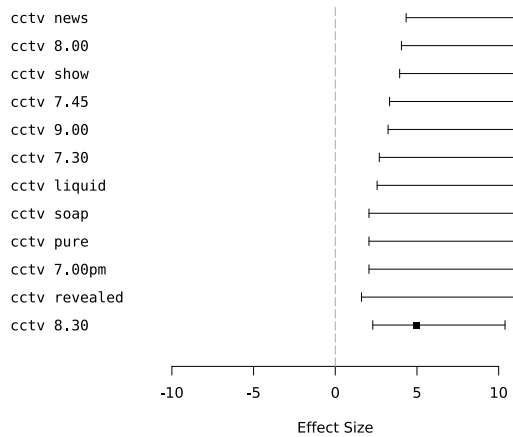
“stop snooping and bugging”. Surveillance scholars have warned that the relationship between the concepts of surveillance and privacy is complex; they function “sometimes in opposition, sometimes as mutually constitutive and occasionally more ironically than we might anticipate” (Ball & Haggerty, 2005, p. 133).

6.3.3 Peak comparisons for *CCTV*

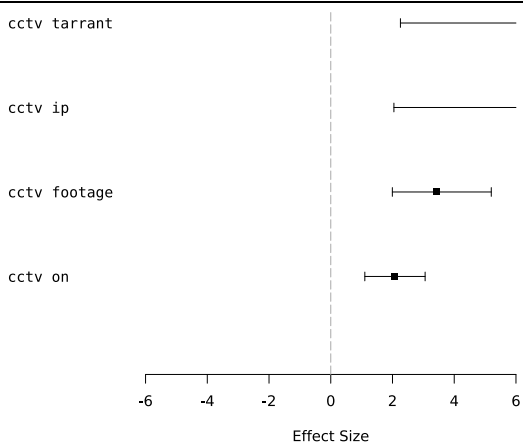
The frequency profile of *CCTV* is different from that of the other nodes (see Section 6.3, Figure 6-24). Following relatively low frequencies up until the late 1990s, the five top peaks months occur from 2003 onwards. Due to this concentration of the *CCTV* occurrences towards the later years of the TDA1986–2008, I used the year 1999 as a reference corpus for this comparison. The resulting collocates for each peak are shown in Figure 6-28.



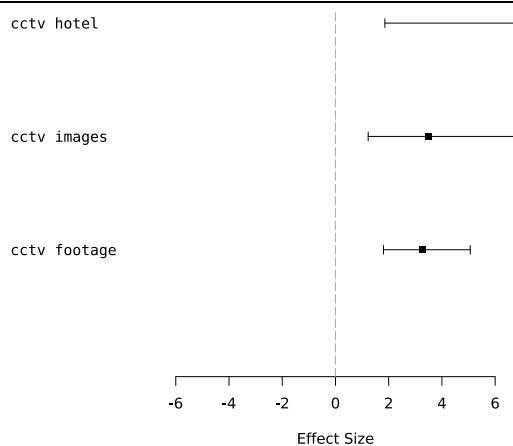
2003_02



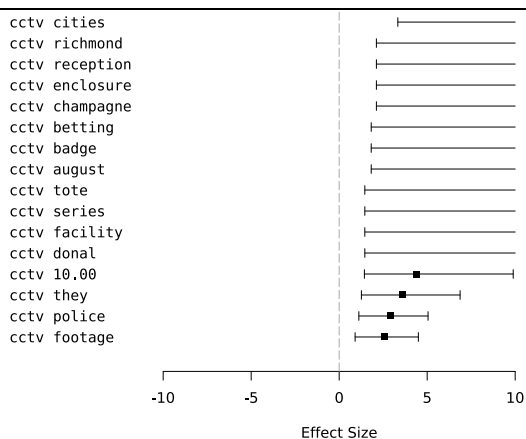
2003_03



2005_08



2007_03



2008_07

Figure 6-28: Co-occurrence comparison of CCTV in each of the top 5 peak months (vs. 1999), ordered chronologically

Most of the collocates from the comparisons with the 2003_02 and 2003_03 peak months are unique. These mainly appear to relate to listings of the *Times* TV programme. “CCTV” was the name of a game show shown on BBC Three from late February to early May 2003. Members of the public were captured on hidden cameras and encouraged via speaker to complete a task (“CCTV - UKGameshows,” n.d.). This, in fact, explains why the programme appears in the March 2003 results of the *CCTV* co-occurrence comparison. The 2005_08 peak appears to be related to a similar programme: TV programmes listing the ITV entertainment show “Tarrant on CCTV”. “Tarrant” is the surname of the presenter, thus explaining the collocates *tarrant* and *on* for this peak in Figure 6-28. Jermyn (2007, pp. 113–114) describes the focus of the programme and points out that the idea of sharing CCTV footage for entertainment is not uncontroversial:

Tarrant on CCTV consists of a variety of CCTV clips presented by the popular TV quiz host Chris Tarrant, obtained from diverse sources, from US police to officers and supermarkets in the UK, and selected for their ‘comedic’ value. [...] but at no point does the programme question the ethics of having CCTV installed in these spaces [...]

The 2005_08 coverage of *CCTV* is not only restricted to the TV programme. *CCTV* also appears in the News section. For example, the reporting after the 21 July London terrorist attack refers to “CCTV footage” being examined. Similarly, there are instances of *CCTV* relating to a potential crime case: references to the death of cricket coach Bob Woolmer are found in the concordances for the 2007_03 peak. The police investigation involves an examination of “hotel CCTV footage”, as illustrated in the snippet in Figure 6-29.

► **Police examining
hotel CCTV footage**

► **TV reports death
by ‘strangulation’**

**Ivo Tennant
Pat Gibson**

Police investigating the death of Bob Woolmer, the Pakistan coach, have seized CCTV footage from the hotel in Jamaica where he was staying. Speculation about the final hours of the former England Test cricketer has reached frenzied levels

Figure 6-29: Beginning of an article from the 2007_03 peak, showing a co-occurrence of *CCTV* with *hotel*

Finally, the collocates revealed in the 2008_07 comparison point to a number of findings. *Police* and *footage* behave in a manner that is similar to what I have shown in the discussion of the year 2008 as a whole (in Section 6.2.3.3) and the hotel CCTV footage in 2007_03. Both collocates are used in reports on (often ongoing) crime cases. However, the collocate *cities* refers to another TV series called “CCTV Cities” (“*CCTV Cities*,” 2016). Many of the unique 2008_07 collocates originate from an ad for a horse race, which took place at the end of July and beginning of August 2008. Accordingly, the ad was repeatedly placed in the lead-up to the event and promotes several features of the race, listing as one of assets “Tote betting facility and CCTV”. This example, therefore, reminds of the housing ads in Section 6.2.3; *CCTV* is mentioned as a reassurance that the premises are safe.

Examples in this section relate to some of the previous findings of CCTV footage being used as evidence without much discussion on the implications of this technology in the near proximity. At the same time, the co-occurrence patterns surrounding the peaks have revealed an aspect of CCTV footage that did not emerge from the broader long-term comparisons of Section 6.2.3. That is, CCTV provides appealing material for popular culture and

entertainment-focused TV shows in particular: the unique collocates from four of the five monthly *CCTV* peaks point to TV shows.

6.3.4 Peaks supporting and complementing long-term trends

This section has examined co-occurrence patterns at the top relative frequency peaks of the three nodes and therefore addressed RQ 3-2. Many co-occurrence results for the monthly peaks appear to echo some of the main findings from the long-term comparisons. This is, unsurprisingly, especially obvious for those peaks that fall into the years from the long-term comparisons (especially 1997 and 2008), where the findings overlap, and, therefore, add further support. In other cases, the peak results from periods that were not directly involved in the long-term comparison resemble general patterns from those found in the full years. For example, the 1993_01 debate on a new civil law of privacy shows similarities with the 1997 discussion of introducing a privacy law and similar uses of CCTV footage are found in reports on the 2007_03 hotel CCTV examination and the general findings on CCTV footage in 2008. Importantly, the bottom-up selection of the peak periods makes it possible to identify patterns that are not confined to the somewhat arbitrarily selected top-down selected periods at the beginning, middle and end of the corpus years (e.g. the “surveillance zone” in the 2007_11 subcorpus).

In summary, while the results from this section reaffirm the findings from Section 6.2, the peaks also highlight locally salient events, provided that these gathered enough recurrent mentions in the given monthly subcorpus. The following section builds on the findings from Sections 6.2 and 6.3 that indicated accumulations of co-occurrences at particular locations in the newspaper.

6.4 Distribution across sections

The discussion of the long-term and peak comparisons has indicated the newspaper section in which a co-occurrence pair occurs plays an important role in the meaning-making of the discourse. Some collocates tend to appear in adverts or TV programmes instead of news in particular temporal periods. These findings suggest that the location of the patterns in the text plays a role in the creation of meaning. In this section I investigate the distribution of the node words across the main newspaper sections for each of the years from the long-term comparisons in Section 6.2. This analysis builds on the argument of previous corpus linguistic research that textual location is important for meaning-making, particularly for the medium of the newspaper (Gupta, 2015; Mahlberg, 2007c; O'Donnell et al., 2012).

My approach shares the interest in meaning, textual location and context with the work of O'Donnell et al. (2012) and Gupta (2015), but differs in its implementation of examining these interactions. I focus on the distribution of the nodes across sections, in a similar manner to previous studies of distributions of particular terms across sections in *The Guardian*: Mahlberg's (2007c) synchronic analysis of the term *sustainable development* in 2002 and Kehoe and Gee's (2009) tracing of *credit crunch* across sections in monthly periods from 2007–2008. Therefore, I consider the phenomenon of textual location on a conceptual level above the focus on articles vs. texts (Gupta, 2015) and the focus on paragraphs (O'Donnell et al., 2012). To further examine the role of the section in the results of the co-occurrence comparison, consider Figure 6-30. It shows the proportion of documents occurring in the main newspaper sections overall (top left panel) and when searching for each of the node words in the TDA interface.¹¹

¹¹ Kehoe and Gee (2009) use a similar visual representation for their analysis of the compound *credit crunch*.

I follow the convention of the interface to use the general term ‘document’, as the overview also contains items like adverts and TV programmes. Each of the bar charts shows the proportions of documents belonging to the main ‘publication sections’ across the years that formed part of the long-term comparison in Section 6.2: 1986, 1997 and 2008. In addition, the data for 2013 is included, as the latest year made available via the expanding TDA online interface at the time of writing (spring 2019). This additional data can give a further indication of whether any tentative trends from previous years are continuing beyond the end of the local full-text dataset, i.e. the TDA1986–2008.

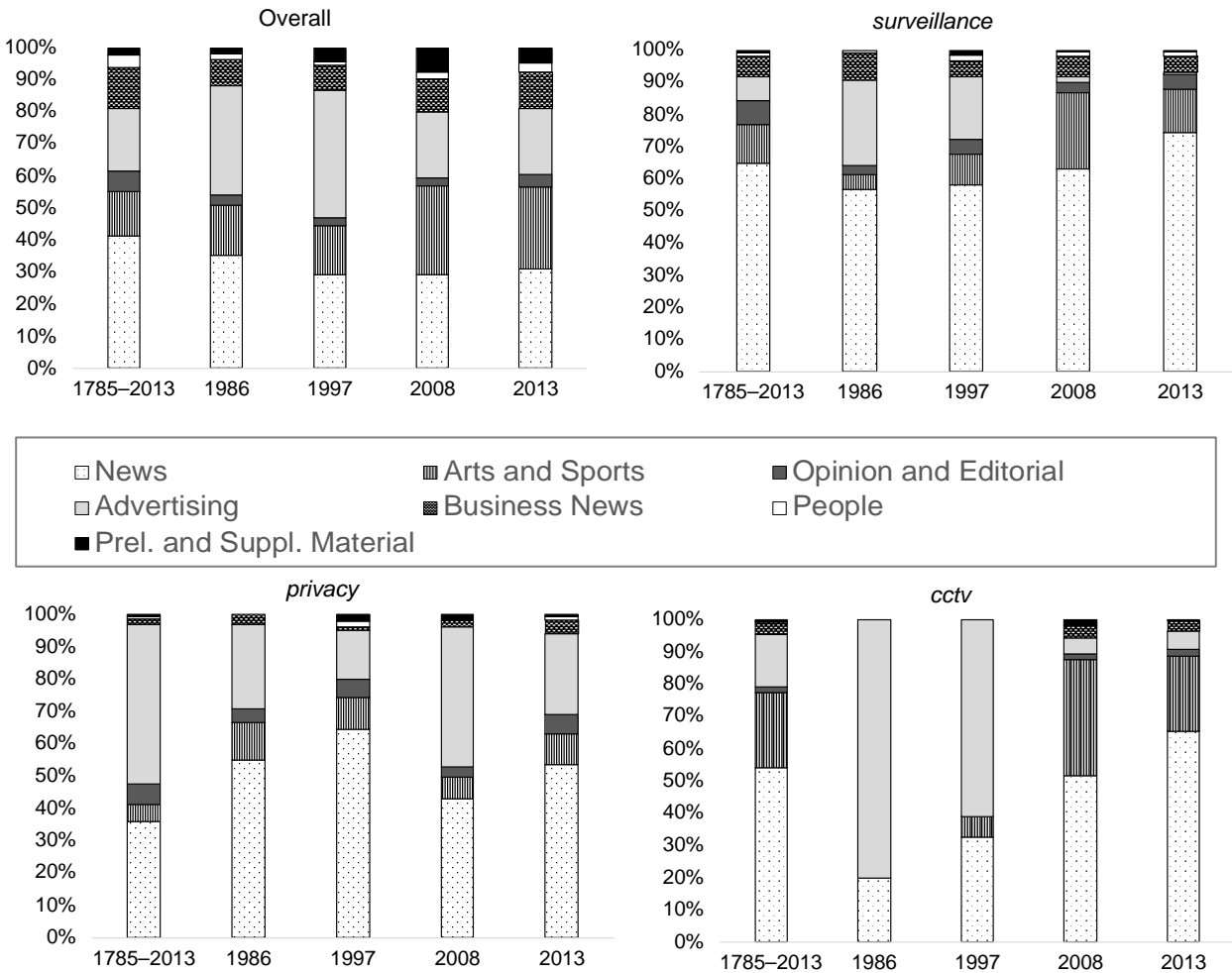


Figure 6-30: Distribution of documents across newspaper sections in the full TDA1785–2013 (first panel) and containing each of the search terms across the newspaper sections per year

The top left panel of Figure 6-30 indicates that the overall proportions of documents from the main categories remain relatively stable across the examined years. “News”, “Advertising” and “Arts and Sports” take up the largest proportion (closely followed by “Business News”). However, the proportion of “Advertising” is higher in the two earlier years, 1986 and 1997, than in the later years as well as than in the overall average. The difference in the proportion of advertising is significant (at $p < 0.0001$) for both 1986 vs. 2008 and 1997 vs. 2008, according to Rayson’s (n.d.) Log-likelihood calculator.¹² This decrease in the proportion of advertising may relate to the changes that the newspaper industry is undergoing, with losses in advertising revenue. However, the advertising documents have a higher chance of being OCR misrecognised so these figures may not be completely accurate. In addition, the plot suggests an increase in the proportion of documents in the “Arts and Sports” section from 1986 to 2013, which is statistically significant ($p < 0.0001$).

The distribution of *surveillance* across the document sections looks different (see the top right panel of Figure 6-30). Accordingly, *surveillance* exhibits a tendency to occur in the “News” section. As demonstrated by the first panel, this trend is not caused by the overall proportion of the “News” documents in the TDA (41%). Diachronically, the proportion of “News” documents further appears to increase: the change from 1986 to 2013 is significant (though only at $p < 0.01$). By contrast, the proportion of advertising documents decreases sharply from 1986/1997 to 2008/2013. This decrease is statistically significant from 1986 to 2013, $p < 0.0001$.

The third panel in Figure 6-30 shows that a relatively large proportion of documents containing *privacy* appear in the “News” section. The proportion is similar but not quite as high

¹² All significance test results in this section were found with this calculator.

as it is for *surveillance*. The year 1997 shows the highest proportion of “News” documents, which matches the findings from Sections 6.2.2 and 6.3.2 of 1997 displaying a spur in discussions of privacy in the context of the press and paparazzi, following the death of Princess Diana. The difference of the proportion in News articles between 1997 and the overall average for *privacy* is significant ($p < 0.0001$). The *privacy* plot also differs from the others in that the proportion of documents in the “Advertising” section does not decrease towards the later years. In 2008, 43% of the documents containing *privacy* are categorised as advertising. While this is a large proportion, it is still below the overall average of *privacy* documents across the TDA (49%; see the first bar). These findings suggest that almost every second document containing the node *privacy* in the TDA is found in advertising. Accordingly, it is not surprising that this distribution was reflected in the co-occurrence patterns with the high proportion of 2008 *privacy* ads featuring, among others, collocates from the terms and conditions of advertised services (see Section 6.2.2).

The final panel illustrates that, for *CCTV*, adverts take up a large proportion of the earlier years, as pointed out in the long-term comparison (Section 6.2.3). However, the contrast with the overall average shown in the first bar indicates that this distribution is not representative. The larger proportion of “News” and “Arts and Sports” found in 2008 and 2013 is more typical. Due to the low frequency of *CCTV* itself in 1986 and 2008 (as the diachronic frequency profile in Section 6.3 has shown), the first two bars are based on much lower frequencies than the other panels. The year 1986 in particular is based on only five documents and the year 1997 is not particularly frequent either compared to the later years and the number of documents containing the other nodes, as shown in Table 6-1. The early distributions of *CCTV* therefore reflect the recency of the technology and are not representative. Conversely, the later years appear much more similar to the overall average, as the bulk of the occurrences belong to this period. This

tendency towards more news and feature articles in 2008 is also reflected in examples of *CCTV* collocates like *footage* and *is* in Section 6.2.3, as distinct from the advert-focused collocates of the earlier years.

Table 6-1: Number of documents in the TDA (overall and containing each of the node words)

	Overall	<i>surveillance</i>	<i>privacy</i>	<i>CCTV</i>
1785–2013	12,252,054	17,922	37,426	4,528
1986	129,760	257	200	5
1997	186,876	324	560	77
2008	167,814	468	955	327
2013	116,394	407	608	276

In summary, this section has addressed RQ 3-3 by examining the frequency distributions of the nodes across newspaper sections. The results from the TDA Online interface revealed that both *surveillance* and *CCTV* occur in a larger proportion of news articles in 2008 than in the earlier years, whereas their proportion in advertising decrease. This observation suggests that both nodes increasingly become part of the main news discourse over course of the corpus. This trend continues beyond the corpus to 2013. *Privacy* is distinct in that its baseline distribution contains a large proportion of adverts. Unlike the other two nodes, the news proportion of *privacy* does not increase steadily, but peaks in 1997. This observation matches the locally salient patterns of the privacy law discussions found for 1997 in Sections 6.2 and 6.3. In addition, it suggests that while *privacy* can contribute to the surveillance discourse, its patterns extend beyond that discourse.

6.5 Case study: The UK Identity Cards Act 2006

Unlike the previous sections, which were based on general node words presumed to relate to surveillance discourse, this section starts from an externally defined topic that relates to a time

frame. The focus is on the coverage of a UK Act of Parliament which received a lot of attention by both the media and academics (Barnard-Wills, 2012; Whitley, 2009/2011; Whitley, Hosein, Angell, & Davies, 2007). In the early 2000s, the government presented a new plan for a national identity card scheme which was finally passed as the Identity Cards Act 2006. The formation of this law was not without controversy and the law was finally repealed by the new government in 2010. Before the initial legislation was passed, it was hotly debated, with issues of privacy and government control at its core. It was an important event in the media representation of surveillance in the UK. One of the academic experts on the UK identity card policies argues that

[...] there has been a marked shift in media coverage of issues of privacy and surveillance. This shift is contemporaneous with the UK Government's plans to introduce biometric identity cards in the UK and suggests that the early debates about the scheme has placed the broader issues of civil liberties onto the mainstream media agenda. (Whitley, 2009/2011, p. 149)

Lyon (2009) shows how closely the issuing and use of national identity cards and passports is tied to questions of social control and state surveillance. He explains that modern identity documents differ from their historical predecessors in two main ways. First, they are characterised by new security features, often including biometric data like finger prints. Secondly, modern identity cards and passports are linked to databases, which enable the government to collect further data.

The cancellation of the act “allow[s] us to perform something of a post-mortem analysis across the life-time of the scheme”, as Barnard-Wills (2012, p. 38) suggests in his governmentality-focused analysis of the identity card discourse. The build-up to the cancellation of the scheme is also of interest for the present work, because this must have occurred conjointly with a discursive shift surrounding identity cards. The Identity Cards Act 2006 originated from a proposal by the Home Secretary of the time to introduce “entitlement

cards” (Whitley, 2009/2011, p. 135). Table 6-2 shows important events in the development of the scheme, from the proposal to a published bill that underwent readings in the House of Commons and the House of Lords and was passed in 2006. As indicated in this overview, the debate of the bill among politicians was also influenced by academic and media reports. A group of academics at the London School of Economics (LSE) formed the Identity Project in order to study the proposed identity card scheme and to inform a public debate. The group published an interim report in March 2005, at the time of the Second Reading in the House of Lords. The bill was then suspended pending the results of the May general election, in which the Labour government was re-elected. In June, the LSE team published its main report (LSE Identity Project, 2005), which was not received well by the government. The Identity Cards Act was finally successfully passed with amendments in March 2006.

Table 6-2: Important events in the development of the Identity Cards Scheme (adapted from Whitley, 2009/2011, pp. 135–139)

Year	Event
2002	<ul style="list-style-type: none"> entitlement cards proposed
2003/2004	<ul style="list-style-type: none"> name change from <i>entitlement cards</i> to <i>identity cards</i>
2004	<ul style="list-style-type: none"> <u>November</u>: publication of Identity Cards Bill
2005	<ul style="list-style-type: none"> <u>February</u>: bill passes Third Reading in House of Commons <u>March</u>: bill undergoes Second Reading debate in the House of Lords and publication of the LSE Identity Project interim report; the bill is suspended <u>May</u>: general election <u>June</u>: release of LSE Identity Project main report
2006	<ul style="list-style-type: none"> <u>March</u>: legislation passed with amendments <u>April</u>: creation of new Identity and Passport Service (in place of the Passport Agency) <u>July</u>: the <i>Sunday Times</i> publishes leaked emails on the front page, titled “ID cards doomed” <u>December</u>: release of new Strategic Action Plan (iris biometrics no longer mandatory & reusing existing government databases instead of creating a new one)
2008	<ul style="list-style-type: none"> <u>March</u>: redesign of the scheme with new delivery plan

Table 6-2 continues beyond the passing of the Identity Cards Act in March 2006. A new “Strategic Action Plan” with changes to the scheme was published at the end of the year,

following leaked emails from senior members of the Office of Government Commerce and the Identity and Passport Service (see Whitley, 2009/2011, p. 138) that were published in *The Sunday Times*. As these emails presented insiders' doubts about the feasibility of the scheme, they had an impact on the overall discourse and the new Home Secretary arranged a review of the entire scheme. Even though the leak does not emerge as an obvious factor from the TDA data presented here, this incident suggests that, like its Sunday paper, *The Times* could play an active role in the debate of the identity cards scheme.

The general timeline of events is reflected in the coverage by the *Times*, including the frequency development of the bigrams *id cards*, *identity cards* and *entitlement cards*, as shown in Figure 6-31. The raw frequency of all three bigrams rises across the early 2000s, peaking in 2005, coinciding with the readings in the House of Commons and the House of Lords. Similarly, the bigram frequencies mirror the name change of the scheme: the original name, *entitlement cards*, only occurs in the 2001 subcorpus (once), 2002 (n = 18) and 2003 (n = 10). Example (25) shows the one occurrence of *entitlement cards* in 2001 (October). This is an early mention of the term by the Home Secretary. The term *identity cards* is already given as a viable alternative. At this point, the option of introducing the cards is presented as requiring careful consideration. However, the article also makes the point that “challenging times [...] demand political action”, pointing to “the events of the past few weeks” (Blunkett, 2001). This statement implies that the September 11 terror attacks require new measures.

(25) And contrary to some commentaries, my position on *identity* or *entitlement cards* remains exactly the same as on 14 September when I was first asked about the issue as Home Secretary. I am persuadable of the case for an entitlement-based card but want to think carefully about it.

(Blunkett, 2001)

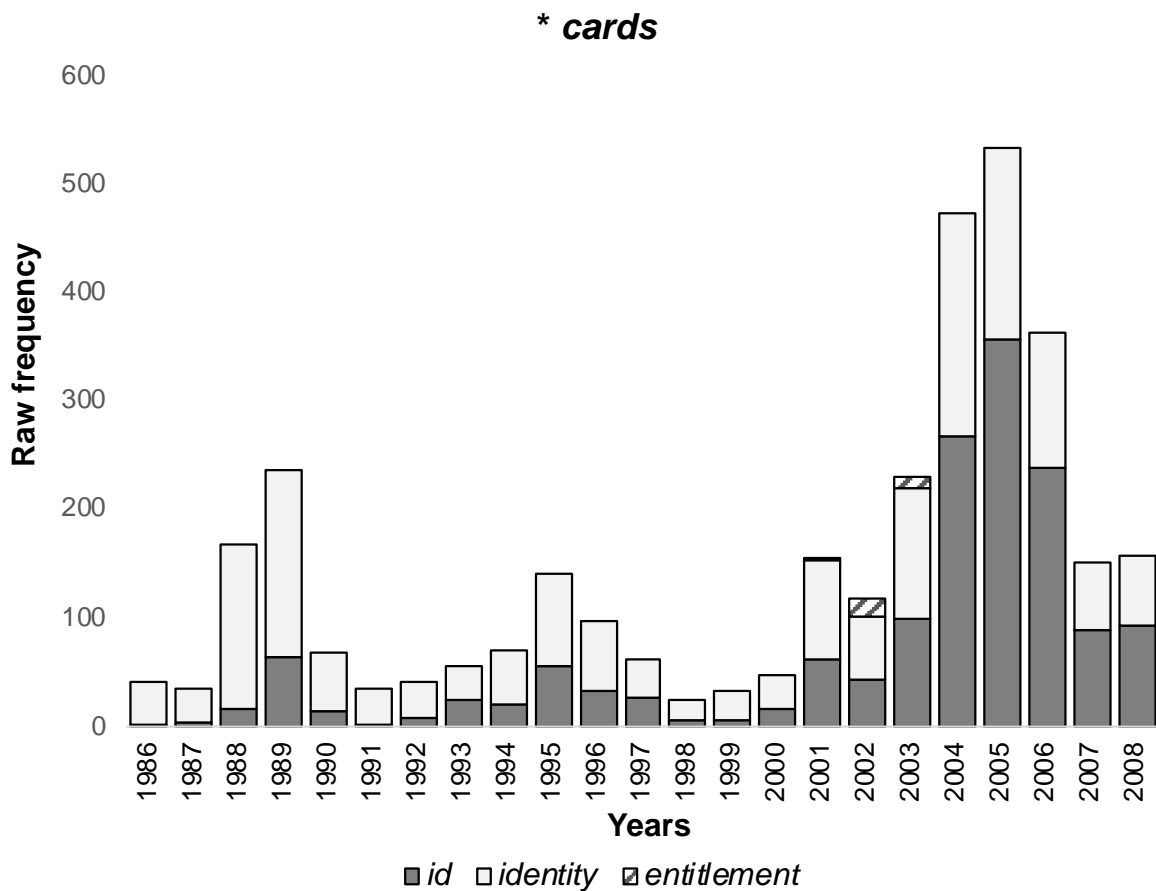


Figure 6-31: Frequency profile of *id/ identity/ entitlement + cards* in the TDA1986–2008

In order to analyse diachronic discursive representation of the Identity Cards Act, I employ the co-occurrence comparison with node words related to identity cards: *identity*, *id*, *entitlement*, *database*, *cards* and the singular form *card*. As these words have various senses beyond the usage in the ID card context, I used a stricter FDR cut-off than I used for the *CorporaCoCo* analysis in Sections 6.2 and 6.3: a rate of 0.01 (which is the recommended default value). For the comparisons in the previous sections I had increased the threshold to 0.1 (thereby allowing for more results), because of the lower frequencies of those nodes.

To investigate the discourse chronologically, I start the discussion with patterns from the first year. Figure 6-31 shows that the bigram *entitlement cards* mainly occurs in 2002 (with a few occurrences appearing in 2003), before the official name changes to *identity cards* (see

Table 6-2). It is therefore not surprising that the co-occurrence comparison reveals significant results for *entitlement* in 2002 collocating with *card* (vs. 2004–2008, all unique apart from 2005) and *cards* (vs. 2006–2008, all unique). Concordance 6-10 displays all instances of this pair in 2002. It presents the *entitlement card (scheme)* as a new proposal, apparent from the quote marks (lines 1, 14, 21) and collocates like *introduction* (line 4), *new* and *mooted* (both line 8), *proposals* (line 18), and the modals on the right side (lines 5, 9, 10, 12, 19–23). These patterns to some extent resemble what we have encountered for the discussion of proposed privacy laws in 1997 (Section 6.2.2) and the 1993_01 peak month (6.3.2).

```
[ 1]           the document as an "entitlement card", which he said would
[ 2] unique personal number. ? Universal entitlement card adapted from credit card
[ 3]           discusses the use of an entitlement card as a proof of
[ 4]           the introduction of an entitlement/identity card scheme ... Your views
[ 5]           New Labour speak I think an entitlement card could offer some important
[ 6]           passport to obtain the entitlement card for £15. If the
[ 7]           already has a form of entitlement or ID card and the
[ 8]           the medical profession. The new entitlement card mooted by the Home
[ 9]           denial of services without an entitlement card? Or should rules be
[10]           driving licence or passport. "The entitlement card scheme could be made
[11]           It is crucial that any entitlement card scheme does not breach
[12]           went ahead with what an entitlement card scheme, it would not
[13]           views in support of an entitlement card scheme have been recorded
[14] secretaries invent euphemisms like "entitlement card", the writing is on
[15]           would pay for the entitlement card through an increase in
[16]           would want to use an entitlement card to replace credit and
[17]           armed with an impressive-looking entitlement card which had been sold
[18]           Kett's proposals for an entitlement card, without real thought about
[19]           considering whether a universal entitlement card would help people to
[20]           billion a year. A universal entitlement card would be a powerful
[21] universal identity card scheme. The "entitlement card" would be needed for
[22]           do not pretend that an entitlement card would be an overwhelming
[23]           scratch. A benefit of an entitlement card would be not having
```

Concordance 6-10: All 23 instances of (*entitlement, card*) in 2002

Example (26) gives the extended context of line 1, illustrating how the term is introduced into the discourse overtly as a direct quote from the Home Secretary. The accompanying indirect speech specifies some of the proclaimed advantages of introducing the scheme. This information functions as a paraphrase in Teubert's (2010) terms, highlighting one aspect of the meaning of the discourse object *entitlement card*. In comparison with Example (25), the quote from the Home Secretary's October 2001 article, Example (26) illustrates that the discourse of

identity or *entitlement card(s)* is already gaining complexity at this early stage. In the extended context of Example (25) the possibility of introducing cards is framed in the context of political action against terrorism. Yet, by the publication of Example (26) in July 2002, the focus has shifted to *improve the delivery of public services* and controlling the benefits for *illegal immigrants*, rather than on fighting terrorism.

(26) David Blunkett, the Home Secretary, described the document as an “entitlement card”, which he said would improve the delivery of public services and make it more difficult for illegal immigrants to claim benefits.

(Ford, 2002)

Another instance of *entitlement card* appearing in quotation marks (from line 14 in Concordance 6-10) is given in Example (27). Here the quote performs a different function, because the choice of the term in itself is the focus. This example portrays the formulation *entitlement card* as a euphemism for a concept that poses a serious danger. Various arguments against the card scheme are put forward, with the negativity implied by the idiom *the writing is on the wall*. The example argues that a natural reaction to the euphemistically titled scheme is to defend *our freedom*. This argument is emphasised with intertextual references to *Agincourt* – the famous speech before the Battle of Agincourt in Shakespeare’s *Henry V* – and the writings of George Orwell.

(27) When Home Secretaries invent euphemisms like “**entitlement** *card*”, the writing is on the wall. So the natural sporting instinct is to haul out every rousing line from *Agincourt* to Orwell, in defence of our freedom to shamble around this precious stone set in a silver sea without carrying some nasty slip of plastic barcoded with our address, health records, fingerprints, financial status and police cautions. We are told that carrying cards would not be compulsory, but the likelihood is that failing to do so would lead to difficulties if you were young, scruffy, dark-skinned or prone to erratic public behaviour.

(Purves, 2002)

It is beyond the scope of this case study to investigate whether this accusation of euphemism in (27) is echoed elsewhere in the early discourse on the identity card scheme and potentially contributes to the eventual name change. However, the co-occurrence comparison shows that in later years *entitlement* co-occurs significantly more often with collocates emphasising its general meaning of having a right to something, such as *sense* and *claim*. Concordance 6-11 positively discusses a “sense of entitlement” among certain groups (“educated women”: line 1; “children”: line 5; “indigenous families”: line 9). The “lack” of entitlement among some groups is a concern (lines 6, 10). On the other hand, a sense of entitlement is criticised if somebody possesses too much of it – “overdeveloped” (line 3), “overblown” (line 4), “disproportionate” (line 8) – or when claimed inappropriately: “Western” (line 2), “insufferable” (line 7). In light of the patterns surrounding “sense of entitlement” and its significance in relation to status and power, the criticism of the card name in (27) may suggest that the cards will cause the opposite effect of what their name claims. Example (27) explicitly suggests that certain people (“young, scruffy, dark-skinned or prone to erratic public behaviour”) may be at risk of being discriminated against in the policing of entitlement cards. This is at odds with the lines in Concordance 6-11, which argue for a need to promote a *sense of entitlement* among “disadvantaged” groups.

```
[ 1]          growing confidence and *sense* of entitlement among these educated women, that
[ 2]          our white, Western *sense* of entitlement, and obligation, to interfere or
[ 3]          having an overdeveloped *sense* of entitlement because it gave the world
[ 4]          with an overblown *sense* of entitlement ("because I'm worth it"), sport
[ 5]          Children have a *sense* of entitlement but there's no legal obligation
[ 6]          She lacks a *sense* of entitlement, even now. Is it fair
[ 7]          I find smokers' *sense* of entitlement fairly insufferable. All that
[ 8]          public a disproportionate *sense* of entitlement in the decision-making process,
[ 9] indigenous families' "legitimate *sense* of entitlement" should be taken into account
[10]          lack a *sense* of cultural entitlement, should be supported in applying
```

Concordance 6-11: Examples of *entitlement* co-occurring with *sense* in 2007 out of 32 total instances

With a focus on change over time, I move on to collocates that are more frequent in the later years compared to the reference year, 2002. In a similar way to the analysis of the shared salient collocates of *surveillance* across the volumes of S&S Corpus (see Section 4.5.1), I trace the diachronic development of collocates related to identity cards. My main concern here is the linear unfolding of a particular story, so I use 2002 as the reference corpus (rather than comparing each year against the whole period, as I did in Chapter 4). The focus is on the most salient collocates that have a link to the debate about the identity cards scheme. Collocates related to, for example, bank, credit and loyalty cards or football (*yellow* and *offsides* in 2008) have been disregarded.

Table 6-3: Selection of co-occurrence pairs significantly different in comparison with 2002

Co-occurrence pair	2003	2004	2005	2006	2007	2008
(card, bill)						
(card, id)						
(card, identity)						
(cards, bill)						
(cards, id)						
(cards, identity)						
(cards, lords)						
(cards, passports)						
(cards, terrorism)						
(id, bill)						
(id, biometric)						
(id, card)						
(id, cards)						
(id, scheme)						
(id, terrorism)						
(identity, bill)						
(identity, cards)						
(identity, passport)						
(identity, service)						

NOTE: light grey = significantly more frequent; dark grey = unique

The results that clearly relate to the introduction of the identity cards scheme are relatively descriptive, including co-occurrence pairs such as (*cards, identity*), (*id, bill*) and (*id, card*), etc. Table 6-3 provides an overview of the most relevant (significantly more frequent) co-occurrence pairs in the later years compared to 2002 (light grey shading). As the two continuous grey rows show, out of this list only the co-occurrence pair (*cards, id*) and its reverse (*id, cards*) are more frequent throughout the entire period than in 2002. The remaining co-occurrence fluctuate more over time.

The largest proportion of the relevant co-occurrence pairs occur in 2005 and 2006 and dark grey cells only appear in these two middle years. They indicate co-occurrences that are unique compared to 2002, suggesting that these collocates present new developments in the discourse. Three of the five unique co-occurrence pairs involve *terrorism*. This may be surprising, since we have seen that the terrorist attacks of September 11, 2001 were part of the context for considerations of introducing identity cards. So, the results of Table 6-3 are not to say that the issue of terrorism is newly raised in 2005/6. It may have been expressed in more indirect patterns in previous years, such as “the events of the past few weeks” in the 2001 quote of the Home Secretary. However, the cumulative picture suggests that it features saliently in the new co-occurrence patterns, (*cards, terrorism*) and (*id, terrorism*), in these middle years. Another unique collocate (from 2006) is *lords*. As Examples (28) and (29) illustrate, this collocate is used in reports on the status of the political debate. This process is depicted as a “battle” in both cases, emphasising the difficulty involved in the formation of the law. The examples also suggest that the patterns of the individual collocates in Table 6-3 are not isolated

from each other. Additional co-occurrence pairs from the table, (*cards*, *terrorism*) in (28)¹³ and (*id*, *biometric*) in (29), are underlined in the examples.

(28) THE long-running battle between Commons and *Lords* over identity **cards** and new anti-terrorism laws continued last night as MPs overruled peers for the third time.
(P. Webster, 2006)

(29) The legislative battle over identity **cards** has ended. The *Lords* and Commons have agreed on a compromise whereby citizens who apply for new or renewed passports before January 1, 2010, will not be required to have an ID card, although their biometric details will go on to a register.
(“The Top Stories,” 2006)

Judging by the results from Table 6-3, co-occurrence change of identity cards from 2002 to 2007/8 provides little novelty. Only three co-occurrence pairs from these years do not appear in the results of the previous years. However, the context of these collocates reflects more changes in the discourse than their surface forms may suggest. The co-occurrence of *identity* with *passport* and *service* both denote an institutional change as a consequence of the Identity Cards Act 2006, referring to the newly established “Identity and Passport Service” (see the timeline in Table 6-2).

Concordance 6-12 shows all 2008 instances of (*id*, *scheme*), the third co-occurrence pair that does not appear in comparisons of the previous years with 2002. The concordance lines highlight that in 2008, so just two years after legislation was passed, the coverage of *The Times* appears to question the future of the scheme. Although *public support* is mentioned several times (lines 3, 4, 25), negative patterns appear to dominate. Also note that in line 25, the verb *claims* suggests that not everyone agrees support is growing. In the extended context we find that the agent claiming the support is a “Home Office-sponsored survey”. Similarly, although

¹³ The tokenisation used for the TDA *CorporaCoCo* comparison separates tokens at hyphens (see Section 3.2.3.2).

line 31 appears positive, the extended context shows that “nobody has given a convincing argument”. In other lines, the scheme is explicitly referred to as “controversial” (2, 16, 27). “Doubts” about the scheme (line 9) and the “uncertainty” of its future (in the extended context of line 29) are expressed.

```
[ 1]         with the ruinously costly national ID card scheme. The data management
[ 2]         way of introducing the controversial ID card scheme, a leaked memo
[ 3]         way ► Public support grows for ID card scheme Richard Ford Home
[ 4]         public support for ID cards ("ID card scheme gathers public support
[ 5]         related to moles) for an ID card scheme. "I regret to
[ 6]         if she would abandon her ID card scheme. The Home Secretary
[ 7]         become ruthless. She wants her ID card scheme and she doesn't
[ 8]         on plans for a national ID card scheme. The MoD said
[ 9]         cut as doubts grow about ID card scheme Murad Ahmed The
[10]         world. The equally ill-conceived ID card scheme is another. When
[11]         the current plans for the ID card scheme, but it sought
[12]         stubborn pursuit of the wasteful ID card scheme. Secondly, simplify taxation
[13]         Howard's support for the Government's ID card scheme before the last
[14]         that is the Government's proposed ID card scheme, which would make
[15]         in their unnecessary and expensive ID card scheme?" Last month the
[16] controversial (and increasingly expensive) ID card scheme. If, however, the
[17]         fiasco fuels calls for £4.7bn ID card scheme to be scrapped
[18]         helped to develop the national ID card scheme, was brought in
[19]         the Government on the national ID card scheme, told the Home
[20]         has fatally wounded the proposed ID card scheme and new projects
[21]         company and confidence in the ID card scheme that it is
[22]         scheme but Secrets of the ID card revealed Remarks: immigration
[23]         Government's proposals for a national ID card scheme. Many were concerned
[24]         or her husband US elections ID cards scheme is pushed back
[25]         that claims public support for ID cards ("ID card scheme gathers
[26]         early assessments of the Government's ID cards scheme began in the
[27]         essments of the Government's controversial ID cards scheme. But Mr. Iust ice
[28]         for vast databases for the ID cards scheme and NHS patients.
[29]         surrounding the future of the ID cards scheme that not only
[30]         enough, the dangerous and illiberal ID cards scheme should be scrapped
[31]         convincing argument for a national ID scheme – and anyway, you will
[32]         Davis voted for a national ID scheme. He may not: have
[33]         various rounds of a Talent ID scheme and was rewarded with
[34]         t otitic d to h id e ich Po t Scheme
```

Concordance 6-12: All 34 instances of *id* co-occurring with *scheme* in 2008

A particular concern is the cost of the scheme: “ruinously costly” (line 1), “expensive” (lines 15, 16) and “wasteful” (line 12). Several lines argue that the scheme is being “pushed back” (line 24), “ill-conceived” (line 10) or suggest it could be “scrapped” altogether/ is “fatally wounded” (lines 17, 20, 30). The immediate context in Concordance 6-12 only shows one instance of civil rights concerns: “dangerous and illiberal” (line 30). However, the extended context of line 11, reproduced in Example (30), shows that the scheme is by now mentioned in

the context of surveillance, although such concerns are played down by politicians' *assurances*. This particular instance also appears in Concordance 6-1 of the *surveillance society* (see Section 6.2.1.3).

- (30) The report, entitled *A Surveillance Society?*, said that it accepted assurances from ministers that surveillance was not part of the current plans for the **ID** card **scheme**, but it sought guarantees that that no expansion would take place without MPs' approval. (Ford, 2008b)

A similar example from 2007 highlighting the proposed national register as an additional surveillance concern besides the cards is given in (31). The particularly critical argument in italics may be explained by the origin of this example (a letter to the editor). To illustrate the opposite perspective, the final example, (32), from 2005 contains a quote from the Home Secretary, who tries to push back allegations that the ID card scheme strengthens the surveillance society by creating more transparency. This attempt to position ID cards as a force against the surveillance society appears to be unsuccessful, as suggested by the reaction of the MPs (laughter).

- (31) The British *government is increasingly becoming a threat to the lives, liberty and property of its citizens*: sexual orientation regulations, environmental policy, ID cards, the smoking ban and the **surveillance** **society**. (Weldon, 2007)

- (32) MPs laughed when Mr Clarke asserted in the Commons that "the ID card system is in fact a bulwark against the **surveillance society**". (Miles, 2005)

The aim of this section was to provide an overview of co-occurrence change in the discourse of identity cards throughout the development of the scheme in response to RQ 3-4. The sociological and surveillance studies literature does feature arguments about a shifting

discourse (see the quote from Whitley, 2009/2011 at the beginning of this section). The corpus linguistic approach presented here provides evidence of changing patterns. This section has demonstrated that the development of the identity cards discussion is reflected in the salient co-occurrences for given years. In addition, concordance lines provide evidence for a negative representation of the ID card scheme in 2008, shortly before the act is repealed in 2010. By tracing the co-occurrences over the period associated with the scheme, I have taken a similar approach to my examination of the shared collocates of *surveillance* in Section 4.5.1 (Table 4-8). The ID scheme has provided a more concrete case study with a timeline to follow across the diachronic structure of the TDA.

6.6 Conclusions

This chapter has shown that the surveillance discourse in *The Times* matures from 1986 to 2008. The chapter has argued that changes in co-occurrence profiles point to changes in the discourse. Accordingly, the representation of surveillance in *The Times* is characterised by both long-term developments and locally salient patterns surrounding the peaks of the analysed nodes. Initially, the surveillance discourse appears near the periphery of the newspaper. The mentions of *surveillance*, *privacy* and *CCTV* in adverts illustrate that they already fulfil social functions in the early part of the discourse as part of company names or the description of services on offer. They only become the main topic of discussion later, for example as news articles start to mention CCTV footage in relation to police investigations. Reflections on the concepts of surveillance and privacy and the use of CCTV technology appear in news articles only towards the latter half of the corpus (RQ 3-1). Influence of individual events is visible from co-occurrences of relative frequency peaks. The privacy law discussion sparked by the publication of photos of Princess Diana have a particularly important impact that is visible in the results for

the year 1997 as a whole and across several peaks. Similarly, *surveillance society* appears in the results for 2008 as a whole and for peaks. Other peaks point to more localised incidents that contribute new dimensions to the surveillance discourse, such as the surveillance zones created in response to disease outbreaks (RQ 3-2). The increasingly “serious” co-occurrence contexts in which the nodes appear is further supported by the expansion of surveillance discourse from the advertising to the main news section of the newspaper seen in Section 6.4 (RQ 3-3). Yet, the distribution trend for *privacy* differs from *surveillance* and *CCTV* for which the proportion of news steadily increases over time. Instead, the news proportion of privacy peaks in 1997 with the accident of Princess Diana, still featuring large proportions of advertising in later years. This diachronic analysis of meaning-making patterns across the sections adds to the literature on meaning and textual location.

Finally, the chapter has illustrated that it is possible to track how particular news story unfolds over the years. It was shown that the co-occurrences reflect the milestones of the debate and indicate an increasingly negative representation towards the end (RQ 3-4). Overall, the chapter has shown that surveillance discourse, and the cultural keyword *surveillance* in particular, has a social relevance throughout the corpus, but is only widely recognised as such towards the end, when the “surveillance society” is actively debated.

7 Conclusions

7.1 Introduction

This chapter concludes the present study with a view to its findings on surveillance discourse, theoretical insights into meaning-making patterns and its methodological contributions. Section 7.2 presents the major findings about surveillance discourse from the three analysis chapters. Section 7.3 draws together the principles of meaning-making that underlie the methodological framework of this thesis and offer a theoretical contribution to corpus linguistic studies of meaning in discourse. In Section 7.4, I discuss the limitations of this study. Section 7.5 suggests directions for future work.

7.2 Major findings

To address the main research question of how surveillance is discursively represented, this thesis has studied meaning-making patterns in three different domains of public discourse. In addition, each analysis chapter focuses on a different thematic component of surveillance discourse: the general meanings of *surveillance* (Chapter 4) and the fundamental discourse coordinates of place (Chapter 5) and time (Chapter 6). This section draws together the major findings from the three analysis chapters.

The study of surveillance discourse in this thesis began with a focus on the meanings of *surveillance* in a corpus of “expert” texts written by surveillance studies scholars. The analysis of this academic discourse was guided by three specific RQs that looked for definitions of *surveillance* (RQ 1-1), consistently salient words across the S&S Corpus (RQ 1-2) and shifts in the co-occurrence patterns of these salient words across the corpus (RQ 1-3). In response to RQ

1-1, the analysis revealed that the scholars describe surveillance as a complex concept. Intertextual patterns showed that several popular academic definitions are referred to repeatedly. One of these is Lyon's (2007, p. 14) definition of surveillance as "the focused, systematic and routine attention to personal details for purposes of influence, management, protection or direction". Despite the popularity, the focus and scope of existing definitions is still open to debate as new contexts and technologies of surveillance emerge. Questions include whether the concept of surveillance can include the watching of non-human subjects and whether it is limited to certain activities (e.g. watching). At the same time, more specific subfields are forming that adapt general definitions for particular application areas. As a result, definitions of *surveillance* are subject to negotiation throughout the corpus, as aspects of definitions are questioned and new definitions are sought for specific research aims.

Chapter 4 identified eight meaning groups of consistently salient words (key keywords; KKWs): (i) theoretical frameworks, (ii) government & the public domain, (iii) time & space, (iv) research & academic writing, (v) social actions & actors, (vi) monitoring & identification, (vii) technology, and (viii) notable theorists in the field (RQ 1-2). These groups provide a high-level representation of the aboutness of the corpus that has been derived from lexical patterns. The chapter argued that these meaning groups relate to four main themes of the S&S journal that have been previously identified via content analysis: 1. "Classic Surveillance", 2. "Identity-based Surveillance", 3. "Mobility and Stasis" and 4. "Work, Power and Resistance" (Mehrabov, 2015). These themes are more localised, because Mehrabov (2015) has derived these themes based on special issue titles (rather than based on all articles) and his analysis ends with special issues in Volume 9. So, the KKW meaning groups can be viewed as elements of surveillance discourse of the journal as a whole, whereas the four themes indicate trends in the study of surveillance.

The analysis of KKW co-occurrence patterns suggested that the surveillance discourse shifts from “Classic Surveillance” to work on “Identity-based Surveillance” and “Work, Power and Resistance”. The mobility-related theme is more dispersed throughout the publication period from 2002–2015. In addition, the chapter put forward an additional theme, “Mass surveillance” that emerges in the volumes beyond the period covered by Mehrabov (2015). The analysis of these shifts has addressed RQ 1-3 and produced a network of shared collocates of the KKW *surveillance* that both represents the aboutness of the corpus and provides insights into discursive change across the journal volumes.

Chapter 5 studied the representation of surveillance in digital discourse based on blog posts with a thematic focus on the discourse coordinate of place. To do that, the chapter adapted the mediated discourse analysis framework of surveillant landscapes (Jones, 2017) for an analysis of multimodal and textual representations of place in surveillance discourse. The chapter has argued that the principles of mediated discourse analysis and corpus linguistics can be usefully combined. The textual focus of corpus linguistics can complement qualitative insights on the interaction of multimodal elements and mediated actions and expand the scale of the study. The analysis of Chapter 5 therefore took two perspectives. First, a case study of a local shopping centre highlighted the multimodal ways in which surveillant landscapes are represented (RQ 2-1). Second, the analysis studied the textual representation of surveillant landscapes on a larger scale in the Surveillance Blog Corpus (RQ 2-2).

With the help of the multimodal case study, the analysis identified a range of analogue and digital ways in which the centre’s surveillant landscape is multimodally constructed and represented. These elements, or discourses in place, included surveillant features in the built environment such as wide corridors and CCTV cameras, the digital tracking of visitors’ activities and paths via their mobile phones and human surveillance by shop assistants and

police officers. The small print on the centre's websites further contributes to the surveillant landscape. Overall, the analysis suggested that visitors have little control in this surveillant landscape, except where they are encouraged to partake in surveillant activities for security (e.g. reporting suspicious behaviour) or playful reasons (e.g. photographing outfits with an app that provides shopping results) (RQ 2-1).

With the help of semantic co-occurrence comparisons, the textual analysis of surveillant landscapes in the Surveillance Blog Corpus was able to study a large variety of place and mobility references to both physical and virtual places. Importantly, the blog posts were identified lexically, with the S&S Corpus acting as a 'seed corpus'. The corpus linguistic analysis highlighted various factors related to the social relationships constructed by surveillant landscapes. These include questions of freedom, rights, obligations and safety. The context tends to play a critical role in these discussions about attitudes to surveillance: as the chapter showed, only because apparently positive semantic tags are used, the mentioned surveillant practices are not necessarily endorsed. So, the combination of quantitative and qualitative analyses is important (RQ 2-2). In addition, the comparative analysis of the blog posts and academic articles revealed differences related to the affordances and constraints or communicative purposes associated with these discourse domains. Compared to the academic articles, the co-occurrence comparisons demonstrated more time references, which appear to facilitate more immediate responses to current affairs. Similarly, the blog posts appeared to be more concerned with conflicts and financial crime, highlighting similarities to the value of negativity associated with the news domain. Given its heterogeneous background, the Surveillance Blog Corpus additionally contains corporate promotions of measures that will bolster surveillant landscapes.

Chapter 6 focused on the discourse coordinate of time, which closely relates to meaning-making principle (i) that meaning evolves with the discourse. Meaning is as dynamic as the discourse. The chapter argued that the discursive representation of surveillance matures over the period from 1986–2008 in *The Times*. The social relevance of the nodes *surveillance*, *privacy* and *CCTV* is apparent across the long-term development. At the beginning of the period, these nodes are mentioned more frequently as positive elements in advertising on surveillance equipment, or real estate properties. Co-occurrence patterns of *surveillance* and *CCTV* in particular then display a change towards being mentioned as part of a wider range of contexts, in the TV programme and the news. A debate about the advantages and disadvantages only develops toward the end of the corpus, with a salient discussion of the “surveillance society” in 2008 (RQ 3-1). The case of *privacy* is different, because it is most strongly debated in the news 1997, as part of the coverage of Princess Diana’s accident. This event is therefore also reflected in the local peaks around that period. So, the locally salient patterns surrounding relative peaks do not necessarily diverge from long-term patterns, especially when the sampling periods of the two comparisons overlap. Nonetheless, the examination of the peaks has also pointed to more localised stories (RQ 3-2).

The observations from the co-occurrence patterns are supported by the diachronic development of the dispersion of the nodes across newspaper sections. The proportion of adverts mentioning *surveillance* and *CCTV* decreases over the TDA1986–2008 period and beyond, while the news proportion increases. For *privacy*, the news proportion peaks in 1997, in tandem with the co-occurrence and concordance findings. Overall, the analysis of the distribution of nodes across the newspaper sections has supported findings from earlier research arguing that textual location plays an important role in meaning-making (Gupta, 2015; Kehoe & Gee, 2009; Mahlberg, 2007c; O’Donnell et al., 2012). Co-occurrence patterns in advertising

and TV programmes featured prominently in the results. One reason for this salience may be a relatively higher degree of repetition in these categories of the TDA. As films are shown multiple times per year, their titles occur repeatedly in the corpus. Similarly, ads can be placed repeatedly. By contrast, news articles are highly unlikely to be repeated verbatim as a whole although subsections or quotes may be reused. By providing evidence that diachronic changes in textual location affect meaning change, Chapter 6 has contributed new insights to this area of research (RQ 3-3). Unlike the long-term/top-down and peak/bottom-up selection for the previous analysis stages, the final analysis stage of Chapter 6 was based on an external timeline. The case study of the Identity Cards Act 2006 demonstrated how co-occurrence patterns can be analysed to track the unfolding of a particular news story. The co-occurrences results reflected the milestones of the debate and indicated an increasingly negative representation at the end of the period (RQ 3-4), leading up to the abandonment of the ID card scheme. So, Chapter 6 has demonstrated that surveillance discourse is socially relevant across the TDA1986–2008 corpus, but only widely recognised as such towards the end, when the “surveillance society” is actively debated. Patterns surrounding *privacy* have supported this discourse at times, and at other points acted more independently – showing that surveillance and privacy are two related but not always analogous concepts. Methodologically, the analysis of additional semiotic information on the newspaper layout, images of adverts and newspaper sections has supported the textual findings.

Overall, the three analyses of the three different domains of discourse have shown that surveillance is:

- a type of supervision or control that is often associated with watching;
- more complicated than balancing privacy and security;
- increasingly subject to debate;

- a source of empowerment and oppression;
- closely tied to rights and obligations;
- evolving over time and with technological developments; and,
- ubiquitous; both in terms of place and the contexts in which it plays a role.

In this sense, *surveillance* has a similarly elusive quality to words and discourse concepts that have been termed cultural keywords, including *sustainable development* (Mahlberg, 2007c; see Sections 2.4, 2.5.3 and 4.2.2) or *globalisation* (Teubert & Čermáková, 2004; although they do not use the term “cultural keyword”; see Section 2.4.5).

7.3 Principles of meaning-making in discourse

This thesis has argued meaning has to be studied in comparison. Because meaning is context-bound, I have focused on a particular discourse object, surveillance. This section reflects on the three ‘principles of meaning-making’ that this thesis has put together to understand surveillance: (i) meaning evolves with the discourse, (ii) meaning emerges via comparison and (iii) meaning takes shape in co-occurrence patterns.

(i) Meaning evolves with the discourse

If society keeps adapting and evolving, how can discourse not change? The present work has taken the view that meaning arises from social interaction in discourse. This is a dynamic process. One dimension along which meaning evolves is diachronic (see Teubert, 2010). Change over time has been implicitly studied across the volumes of the S&S journal in Chapter 4, indicating developments in the theoretical approaches to surveillance and technological advances. Chapter 6 has explicitly focused on time, making use of the temporal structure of the

TDA. Another dimension along which meaning changes is synchronic: discourse domains and, within them, textual location. This argument is inherent to the overall methodological framework that is based on corpora from different discourse domains. In Chapter 3, I have explained that within the corpora, it is possible to differentiate between levels of meaning-making that interact in their contribution to the discourse. The S&S journal can be seen as a whole, but information about how lexical patterns behave across volumes, issues, articles adds another level of understanding to the analysis. Similarly, different blogs and their individual posts bring their unique perspectives on surveillance. In different newspaper sections, surveillance nodes develop specific functions. Textual location therefore emerged as an important aspect of meaning-making in Chapter 6 in particular, but was already inherent in the methodological framework for the corpus compilation in Chapter 3.

(ii) Meaning emerges via comparison

This thesis argued that it is only possible to realise which elements of a corpus are salient in comparison with other corpora where these elements are not as striking or common. The analysis has applied comparisons within and across the corpora compiled for this thesis. The main research question of the present study is concerned with the representation of surveillance in public discourse. In order to take any practical steps in this analysis, it was necessary to operationalise the concept of public discourse and to break it down into particular domains. I have focused on three broad domains: academic discourse, digital discourse (represented by the blogosphere) and news discourse. The three domains revealed their own individual concerns with surveillance, overlapping only in certain aspects. Examples of these overlapping concerns include areas like the security of sports events, which is a topic in a special issue of the S&S Corpus, and received criticism by one of the blog posts analysed in Chapter 5 and the debate

on identity cards. The latter provides a clear case in point of how the different discourse domains can influence each other, because an academic report (LSE Identity Project, 2005) had a salient influence on the discussion of identity cards in parliament and in the media.

(iii) Meaning takes shape in co-occurrence patterns

Co-occurrence patterns are at the core of the analysis of meaning in this thesis. Following meaning-making principle (ii), the co-occurrences have been analysed comparatively, while paying attention to the context of examples in the three discourse domains. Co-occurrence patterns have not been considered in isolation: the discussion of the network of shared collocates in Chapter 4 has supported recent arguments that meaningful links exist between collocates in discourse (see Brezina et al., 2015). Similarly, the present work has built on research that takes co-occurrence beyond its lexical role by extending the analysis to semantic comparisons in Chapter 5 (cf. association measure-based semantic tag collocates; Prentice et al., 2012). In addition, the thesis has put forward ways of examining the “co-occurrence” between textual patterns and elements of other semiotic modes (Chapters 5 and 6).

7.4 Limitations

The co-occurrence comparison methodology is novel. Since the present study presents the first large-scale application, the method is still in its early stages of development. Some limitations have been noted in Section 3.3.3 relating to the current restriction of *CorporaCoCo* to the analysis of node words rather than multi-word units. In addition, the approach encourages a focus on differences rather than similarities in the data. I have attempted to counterbalance that tendency in this thesis by analysing results in context. In Chapter 4, I have further identified KKWs and shared collocates, which both helped to appreciate consistency across the

subcorpora. Due to reasons of space, it was not possible to integrate similar steps into the analysis procedures of Chapters 5 and 6. However, the analysis of concordance lines and examples in those chapters along with the additional multimodal material has hopefully provided extensive insights into discursive patterns beyond significantly different collocates.

The three corpora studied in this thesis and the discourse domains that they are representing only cover a fraction of the discourse on surveillance. Studying additional domains was beyond the scope of this study. Among the corpora analysed, the S&S Corpus presents the largest coverage compared to the full extent of the discourse domain. This corpus misses all the books published on the subject, articles that have appeared in related journals and any S&S articles from winter 2015 onwards. Nevertheless, it provides a rather good approximation of the English-language academic discourse on surveillance.

By contrast, both the Surveillance Blog Corpus and the TDA1986–2008 can only provide a small fraction of the texts written about surveillance in their domains. Therefore, the compilation of these corpora involved more challenging decisions. For the design of the present study it was desirable to approach the compilation of the different corpora following a variety of criteria. The S&S Corpus was clearly based on external criteria, so internal criteria were chosen for the blog posts. This way of collecting the texts made a useful link between the lexical patterns in the seed corpus, the S&S Corpus which was by definition “about surveillance”. However, the variety of texts on the internet begs the question as to what counts as a “genuine” or “valuable” blog post for a corpus linguistic study of discourse. The Surveillance Blog Corpus has been “cleaned” of obvious duplicates and noise that appeared to be spam (see Section 3.2.2) because it was not straightforward to make value judgements about which texts to include. Therefore, the posts present a large range of apparently personal, ideological and corporate interests, as well as a range of copy-pasted chunks from other sources. Smaller scale studies

with a more qualitative focus tend to use strict external criteria in their selection of blogs (e.g. Myers, 2010; Zou & Hyland, 2019; Hoffmann, 2012; Puschmann, 2010). While such selective approaches ensure a certain quality of a controlled dataset, they can only represent very small sections of the blogosphere, (e.g. 31,000 words from 30 posts in Zou & Hyland, 2019's study). Larger scale studies (e.g. Biber & Egbert, 2016; Titak & Roberson, 2013), cannot retrieve the same level of detail about the history and context of individual blogs. The present study has attempted to find a middle ground. Rather than follow external selection criteria, Chapter 5 has aimed to approach meaning-making patterns in blog discourse from a lexically-driven perspective, using the S&S Corpus as a 'seed corpus'.

As a mega-size corpus, the TDA1986–2008 has provided a valuable data source for the present study. Its powerful size presents challenges that make the processing less flexible than the analysis of smaller corpora. Once processed, the format of this dataset can only be changed with considerable investment of time (and possibly cost). So, it was not possible to semantically tag this corpus as was done for the two much smaller specialised corpora analysed in this study. A more impactful disadvantage for the present study is that the processed portion of the corpus ends in 2008, well before the 2013 leak of NSA documents (cf. Sections 2.2; 4.4.2.3), which was widely covered by news media and is likely to also have affected patterns of *surveillance* in *The Times*, at least in that year. Nevertheless, it is hoped that the analysis of the “recent history” of *surveillance* in this newspaper is still valuable. A wider selection of newspapers would have provided more variation, but a comparable corpus was not available. Moreover, the focus on a single newspaper also allowed for a more detailed analysis of that publication (cf. Zinn & McDonald, 2017).

7.5 Directions for future work

The research reported has opened up new lines of inquiry related to multiple application areas. An obvious area for which this work has implications is the study of meaning in discourse along the lines of meaning-making principles this study has put together. In this section I offer further methodological implications of the present work.

The main practical contributions of this thesis lie in its comparative analysis framework, which is both developed in the selection and compilation of corpora and their analysis via co-occurrence comparisons. The present work has made a case for compiling corpora from different domains of the public discourse to facilitate comparison. To do so, the study has followed different techniques of corpus compilation based on meaning-making levels of the given text type with text-internal and -external criteria. The study has introduced the notion of the ‘seed corpus’ for using an existing specialised corpus to create a corpus from another domain with a shared lexical starting point. More work can be done to explore methods for text-internal criteria of compilation that link the previous research on relevance (see Gabrielatos, 2007) with theoretical work on meaning in discourse.

This thesis is the first large-scale application of the co-occurrence comparison method. By virtue of testing out early versions of *CorporaCoCo*, the present work has both benefitted from being part of the development of that method and had to cope with the situation that the methodological framework was still subject to change. Areas for future development on the co-occurrence comparison method include ways of taking dispersion into account, for example via mixed-effects models (a start was made in Wiegand, Hennessey, Tench, & Mahlberg, 2017a), and the possibility to deal with multi-word units. Currently the challenges are to determine the span for surface co-occurrence for discontinuous multiword units and, in the case of continuous ones, to count the hits and misses accurately across the corpus once the basic definition of a

token has changed, because many of them will now be part of multiword units (see Section 3.3.3). Future research should also continue looking for ways of combining quantitative results from the *CorporaCoCo* package with qualitative analysis.

There is scope for developing the corpus linguistic concept of co-occurrence beyond the practical, applied sense of counting words in order to recognise its potential for meaning-making. Once we recognise that, the basic idea that meaningful effects from different elements in the environment (text) come together to create meaning, can create links to other disciplines. The present study has demonstrated links with mediated discourse analysis. This work on the surveillant landscape framework (Jones, 2017) can be further extended to surveillance studies, to be viewed as an “assemblage” (Haggerty & Ericson, 2000) of meanings.

Another area in which more work can be done is the analysis of intertextuality – and this may well contribute to the idea of meaning assemblages. In Chapter 4, I discussed the relevance of intertextual relations to the development of meaning in connection with the DEFINE concordance of *surveillance*. Future work could look into possibilities of combining a concordance analysis with a citation analysis, for example relying on the information of a database like *Google Scholar* to create a corpus (in this case a subset of the S&S Corpus) of articles linked by citations. This might be one step towards the intertextuality and paraphrase methodologies that Teubert (2019) envisions in stand-alone software; it would also overcome the restriction of the analysis in Chapter 4 to rely on a specific lexical marker for identifying definitions or to extend this.

The arguments about meaning change and in particular the methodological framework for diachronic comparisons in Chapter 6 have implications for diachronic and historical linguistics. The method can be applied to many different contexts in order to compare co-occurrence change over time. Ideally, this work will be taken further beyond pairwise

comparisons. Beyond these statistical concerns, there is also scope for more qualitative research. The present work has provided a “recent history” view on surveillance, by contrast to those studies that focus on surveillance-related events and technologies from Snowden onwards. And even though a whole body of work exists in surveillance studies that has also dealt with historical aspects and cultural questions, I would argue that there is potential for more discourse-focused work on the historical development of surveillance as a concept. Two aspects in which I am planning to take this forward include a study of historical classified ads mentioning surveillance (with early beginnings in Wiegand, 2019) and an extension of the identification and documentation discourse that was indicated in the case study of ID cards in Chapter 6. Finally, the analysis can be extended to additional discourse domains. A logical extension would be the analysis of security policies (cf. MacDonald et al., 2013). Another relevant source of data are documents by non-governmental organizations that aim to raise awareness of surveillance and civil liberty issues, such as Liberty or Big Brother Watch.

This large potential for future work underlines the importance of the results and the methodological framework presented in this thesis. The present study makes a major contribution to interdisciplinary research on textual and multimodal approaches of corpus linguistics, discourse analysis and surveillance and security studies. In studying various domains of surveillance discourse, the present study has made a start in rethinking how meaning works in discourse and how it can be studied.

References

- About the Journal. (2019). Retrieved July 2019, from <https://ojs.library.queensu.ca/index.php/surveillance-and-society/about>
- Agar, J. (2005). Identity cards in Britain: Past experience and policy implications. *History & Policy, Policy Paper 33*. Retrieved July 2019, from <http://www.historyandpolicy.org/index.php/policy-papers/papers/identity-cards-in-britain-past-experience-and-policy-implications>
- Albrechtslund, A. (2008). Online social networking as participatory surveillance. *First Monday, 13*(3).
- Alsop, S. (2016). The “humour” element in engineering lectures across cultures: An approach to pragmatic annotation. In M. J. López-Couso, B. Méndez-Naya, P. Núñez-Pertejo, & I. M. Palacios-Martínez (Eds.), *Corpus Linguistics on the Move: Exploring and Understanding English Through Corpora* (pp. 337–361). Leiden: Brill.
- Altmeyer, S., Klein, C., Keller, B., Silver, C. F., Hood, R. W., & Streib, H. (2015). Subjective definitions of spirituality and religion: An exploratory study in Germany and the US. *International Journal of Corpus Linguistics, 20*(4), 526–552.
- Andersen, G. (Ed.). (2012). *Exploring Newspaper Language: Using the Web to Create and Investigate a Large Corpus of Modern Norwegian*. Amsterdam: John Benjamins.
- Andersen, G., & Hofland, K. (2012). Building a large corpus based on newspapers from the web. In G. Andersen (Ed.), *Exploring Newspaper Language: Using the Web to Create and Investigate a Large Corpus of Modern Norwegian* (pp. 1–28). Amsterdam: John Benjamins.
- Androutsopoulos, J. (2010). Localizing the global on the participatory web. In N. Coupland (Ed.), *The Handbook of Language and Globalization* (pp. 201–231). Chichester: John Wiley & Sons.
- Anthony, L. (2016). *AntConc* (Version 3.4.4m) [Computer software]. Available from <http://www.laurenceanthony.net/software/antconc/releases/AntConc344/>
- Anthony, L., & Baker, P. (2015a). *ProtAnt* (Version 1.0) [Computer software]. Available from <http://www.laurenceanthony.net/software/protant/>
- Anthony, L., & Baker, P. (2015b). *ProtAnt*: A tool for analysing the prototypicality of texts. *International Journal of Corpus Linguistics, 20*(3), 273–292.
- App Store. (2017, October 26). Bullring & Grand Central PLUS on the App Store: Version history 2.8.0. Retrieved July 2018, from <https://itunes.apple.com/us/app/bullring-grand-central-plus/id968698594?mt=8>
- Archer, D., Wilson, A., & Rayson, P. (2002). *Introduction to the USAS category system*. Retrieved April 2017, from http://ucrel.lancs.ac.uk/usas/usas_guide.pdf
- Aston, G., & Burnard, L. (1997). *The BNC Handbook: Exploring the British National Corpus with SARA* (1st ed.). Edinburgh: Edinburgh University Press.
- Baker, H., Gregory, I. N., Hartmann, D., & McEnery, T. (2019). Applying Geographical Information Systems to researching historical corpora: Seventeenth-century prostitution. In V. Wiegand & M. Mahlberg (Eds.), *Corpus Linguistics, Context and Culture* (pp. 109–136). Berlin: De Gruyter Mouton.

- Baker, M. (1993). Corpus linguistics and translation studies: Implications and applications. In M. Baker, H. Francis, & E. Tognini-Bonelli (Eds.), *Text and Technology: In Honour of John Sinclair* (pp. 233–250). Amsterdam: John Benjamins.
- Baker, P. (2004a). Querying keywords: Questions of difference, frequency, and sense in keywords analysis. *Journal of English Linguistics*, 32(4), 346–359.
- Baker, P. (2004b). ‘Unnatural Acts’: Discourses of homosexuality within the House of Lords debates on gay male law reform. *Journal of Sociolinguistics*, 8(1), 88–106.
- Baker, P. (2006). *Using Corpora in Discourse Analysis*. London: Continuum.
- Baker, P. (2010). Corpus methods in linguistics. In L. Litosseliti (Ed.), *Research Methods in Linguistics* (pp. 93–113). London: Continuum.
- Baker, P. (2014). “Bad wigs and screaming mimis”: Using corpus-assisted techniques to carry out critical discourse analysis of the representation of trans people in British press. In C. Hart & P. Cap (Eds.), *Contemporary Critical Discourse Studies* (pp. 211–236). London: Bloomsbury.
- Baker, P. (2016). The shapes of collocation. *International Journal of Corpus Linguistics*, 21(2), 139–164.
- Baker, P. (2018). Keywords: Signposts to objectivity? In A. Čermáková & M. Mahlberg (Eds.), *The Corpus Linguistics Discourse: In Honour of Wolfgang Teubert* (pp. 77–94). Amsterdam: John Benjamins.
- Baker, P., Gabrielatos, C., & McEnery, T. (2013). *Discourse Analysis and Media Attitudes: The Representation of Islam in the British Press*. Cambridge: Cambridge University Press.
- Bakir, V. (2010). *Sousveillance, Media and Strategic Political Communication: Iraq, USA, UK*. London: Continuum.
- Ball, K. S., & Haggerty, K. D. (2005). Doing Surveillance Studies. *Surveillance & Society*, 3(2/3), 129–138.
- Balossi, G. (2014). *A Corpus Linguistic Approach to Literary Language and Characterization: Virginia Woolf’s The Waves*. Amsterdam: John Benjamins.
- Bare Bones Software. (2016). *TextWrangler* (Version 5.5.1) [Computer software]. Available from <https://www.barebones.com/support/textwrangler/notes-5.5.html>
- Barnard-Wills, D. (2011). UK news media discourses of surveillance. *The Sociological Quarterly*, 52(4), 548–567.
- Barnard-Wills, D. (2012). *Surveillance and Identity: Discourse, Subjectivity and the State*. Burlington, VT: Ashgate.
- Barni, M., & Bagna, C. (2015). The critical turn in LL: New methodologies and new items in LL. *Linguistic Landscape: An International Journal*, 1(1–2), 6–18.
- Baroni, M., & Bernardini, S. (2004). *BootCaT: Bootstrapping Corpora and Terms from the Web*. In *Proceedings of LREC 2004* (pp. 1313–1316). Retrieved July 2019, from <http://www.lrec-conf.org/proceedings/lrec2004/pdf/509.pdf>
- Bastian, M., Heymann, S., & Jacomy, M. (2009). *Gephi: An open source software for exploring and manipulating networks*. Presented at the International AAAI Conference on Weblogs and Social Media, San Jose, CA.
- Bauman, Z. (2012). *Liquid Modernity*. Cambridge: Polity. (Original work published 2000)
- Bauman, Z., & Lyon, D. (2013). *Liquid Surveillance: A Conversation*. Cambridge: Polity.
- Bednarek, M., & Caple, H. (2014). Why do news values matter? Towards a new methodological framework for analysing news discourse in Critical Discourse Analysis and beyond. *Discourse & Society*, 25(2), 135–158.
- Bednarek, M., & Caple, H. (2017). *The Discourse of News Values: How News Organizations Create Newsworthiness*. Oxford: Oxford University Press.

- Bell, A. (1991). *The Language of News Media*. Oxford: Blackwell.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1), 289–300.
- Bennett, C. J., Haggerty, K. D., Lyon, D., & Steeves, V. (Eds.). (2014). *Transparent Lives: Surveillance in Canada*. Edmonton: Athabasca University Press.
- Benoit, K., & Nulty, P. (2016). *quanteda: Quantitative Analysis of Textual Data*. Retrieved January 2017, from <https://github.com/kbenoit/quanteda>
- Bevitori, C. (2010). *Representations of Climate Change – News and Opinion Discourse in UK and US Quality Press: A Corpus-Assisted Discourse Study*. Bologna: Bononia University Press.
- Beynon-Davies, P. (2011). The UK national identity card. *Journal of Information Technology Teaching Cases*, 1(1), 12–21.
- Biber, D., Connor, U., & Upton, T. A. (2007). *Discourse on the Move: Using Corpus Analysis to Describe Discourse Structure*. Amsterdam: John Benjamins.
- Biber, D., & Conrad, S. (2009). *Register, Genre, and Style*. Cambridge: Cambridge University Press.
- Biber, D., & Egbert, J. (2016). Register variation on the searchable web: A multi-dimensional analysis. *Journal of English Linguistics*, 44(2), 95–137.
- Bingham, A. (2007). ‘Drinking in the Last Chance Saloon’: The British press and the crisis of self-regulation, 1989–95. *Media History*, 13(1), 79–92.
- Blood, R. (2002). *The Weblog Handbook: Practical Advice on Creating and Maintaining Your Blog*. Cambridge, MA: Perseus.
- Bondi, M. (2009). In the wake of the terror: Phraseological tools of time setting in the narrative of history. In M. Charles, D. Pecorari, & S. Hunston (Eds.), *Academic Writing: At the Interface of Corpus and Discourse* (pp. 73–90). London: Continuum.
- Bootcat:release_notes:0.71 [Docs]. (n.d.). Retrieved April 2019, from [http://docs.sslmit.unibo.it/doku.php?id=bootcat:release_notes:0.71&s\[\]=utf&s\[\]=8](http://docs.sslmit.unibo.it/doku.php?id=bootcat:release_notes:0.71&s[]=utf&s[]=8)
- Branum, J., & Charteris-Black, J. (2015). The Edward Snowden affair: A corpus study of the British press. *Discourse & Communication*, 9(2), 199–220.
- Brezina, V., McEnery, T., & Wattam, S. (2015). Collocations in context: A new perspective on collocation networks. *International Journal of Corpus Linguistics*, 20(2), 139–173.
- Brown, G., & Yule, G. (1983). *Discourse Analysis*. Cambridge: Cambridge University Press.
- Bruns, A. (2017). Blog. In B. S. Turner (Ed.), *The Wiley-Blackwell Encyclopedia of Social Theory* (pp. 1–3). Chichester: John Wiley & Sons.
- Bruthiaux, P. (1996). *The Discourse of Classified Advertising: Exploring the Nature of Linguistic Simplicity*. Oxford: Oxford University Press.
- Bull Ring, Birmingham. (2018). In *Wikipedia*. Retrieved July 2018, from https://en.wikipedia.org/w/index.php?title=Bull_Ring,_Birmingham&oldid=851675829
- Bullring & Grand Central Birmingham. (2017). *Bullring & Grand Central Style Seeker Ad*. Retrieved July 2018, from https://www.youtube.com/watch?time_continue=59&v=Ki4eXx6_4UQ
- Bullring & Grand Central—Plus. (n.d.). Retrieved June 2018, from <https://www.bullring.co.uk/plus>
- Bullring & Grand Central—Privacy Policy. (n.d.). Retrieved June 2018, from <https://www.bullring.co.uk/privacy-policy>

- Busse, B. (2019). Patterns of discursive urban place-making in Brooklyn, New York. In V. Wiegand & M. Mahlberg (Eds.), *Corpus Linguistics, Context and Culture* (pp. 13–41). Berlin: De Gruyter.
- CCTV - UKGameshows. (n.d.). Retrieved August 2018, from <http://www.ukgameshows.com/ukgs/CCTV>
- CCTV Cities. (2016). In *Wikipedia*. Retrieved March 2019, from https://en.wikipedia.org/w/index.php?title=CCTV_Cities&oldid=754722864
- Čermáková, A., & Mahlberg, M. (2018). Introduction. In A. Čermáková & M. Mahlberg (Eds.), *The Corpus Linguistics Discourse: In Honour of Wolfgang Teubert* (pp. 2–8). Amsterdam: John Benjamins.
- Ceyhan, A. (2008). Technologization of security: Management of uncertainty and risk in the age of biometrics. *Surveillance & Society*, 5(2), 102–123.
- Charteris-Black, J. (2014). *Analysing Political Speeches: Rhetoric, Discourse and Metaphor*. Houndmills: Palgrave Macmillan.
- Chen, M., & Flowerdew, J. (2018). A critical review of research and practice in data-driven learning (DDL) in the academic writing classroom. *International Journal of Corpus Linguistics*, 23(3), 335–369.
- Cheng, W. (2006). Describing the extended meanings of lexical cohesion in a corpus of SARS spoken discourse. *International Journal of Corpus Linguistics*, 11(3), 325–344.
- Cheng, W. (2012). *Exploring Corpus Linguistics: Language in Action*. Abingdon: Routledge.
- Cheng, W., & Lam, P. W. Y. (2010). Media discourses in Hong Kong: Change in representation of human rights. *Text & Talk*, 30(5), 507–527.
- Cheng, W., & Lam, P. W. Y. (2013). Western perceptions of Hong Kong ten years on: A corpus-driven critical discourse study. *Applied Linguistics*, 34(2), 173–190.
- Clarke, I., & Grieve, Dr. J. (2017). *Dimensions of Abusive Language on Twitter*. In *Proceedings of the First Workshop on Abusive Language Online* (pp. 1–10). Vancouver: ACL. Retrieved November, 2017 from <https://doi.org/10.18653/v1/W17-3001>
- CLAWS Input / Output Format Guidelines. (n.d.). Retrieved April 2017, from <http://ucrel.lancs.ac.uk/claws/format.html>
- Conboy, M. (2010). *The Language of Newspapers: Socio-Historical Perspectives*. London: Continuum.
- Culpeper, J. (2001). *Language and Characterisation: People in Plays and Other Texts*. Harlow: Longman.
- Culpeper, J. (2009). Keyness: Words, parts-of-speech and semantic categories in the character-talk of Shakespeare's *Romeo and Juliet*. *International Journal of Corpus Linguistics*, 14(1), 29–59.
- Davies, M. (2008). The Corpus of Contemporary American English (COCA): 560 million words, 1990-present [Online corpus tool]. Available from <https://www.english-corpora.org/coca/>
- Davies, M. (2010). More than a peephole: Using large and diverse online corpora. *International Journal of Corpus Linguistics*, 15(3), 412–418.
- De Smedt, K. (2012). Ash compound frenzy: A case study in the Norwegian Newspaper Corpus. In G. Andersen (Ed.), *Exploring Newspaper Language: Using the Web to Create and Investigate a Large Corpus of Modern Norwegian* (pp. 241–255). Amsterdam: John Benjamins.
- Deleuze, G., & Guattari, F. (2005). *A Thousand Plateaus: Capitalism and Schizophrenia* (B. Massumi, Trans.). Minneapolis, MN: University of Minnesota Press. (Original work published 1987)

- Draper, N. (2012). Group power: Discourses of consumer power and surveillance in group buying websites. *Surveillance & Society*, 9(4), 394–407.
- Durrant, P. (2009). Investigating the viability of a collocation list for students of English for academic purposes. *English for Specific Purposes*, 28(3), 157–169.
- Editorial Policies. (2014). Retrieved February 2016, from <http://library.queensu.ca/ojs/index.php/surveillance-and-society/about/editorialPolicies#focusAndScope>
- Egbert, J. (2015). Publication type and discipline variation in published academic writing: Investigating statistical interaction in corpus data. *International Journal of Corpus Linguistics*, 20(1), 1–29.
- Elgesem, D., Feinerer, I., & Steskal, L. (2016). Bloggers' responses to the Snowden affair: Combining automated and manual methods in the analysis of news blogging. *Computer Supported Cooperative Work (CSCW)*, 25(2–3), 167–191.
- Elgesem, D., & Salway, A. (2015). *Traitor, whistleblower or hero? Moral evaluations of the Snowden-affair in the blogosphere*. Presented at the Corpus Linguistics 2015 Conference, Lancaster.
- ESRC Centre for Corpus Approaches to Social Science (CASS). (n.d.). BNC2014. Retrieved July 2019, from <http://cass.lancs.ac.uk/bnc2014/>
- Evert, S. (2008). Corpora and collocations. In A. Lüdeling & M. Kytö (Eds.), *Corpus Linguistics: An International Handbook* (pp. 1212–1248). Berlin: Walter de Gruyter.
- Flowerdew, J. (1991). Pragmatic modifications on the 'representative' speech act of defining. *Journal of Pragmatics*, 15(3), 253–264.
- Flowerdew, J. (1992). Definitions in Science Lectures. *Applied Linguistics*, 13(2), 202–221.
- Franklin, B. (Ed.). (2008). *Pulling Newspapers Apart: Analysing Print Journalism*. Abingdon: Routledge.
- Gabrielatos, C. (2007). Selecting query terms to build a specialised corpus from a restricted-access database. *ICAME Journal*, 31, 5–43.
- Gabrielatos, C. (2018). Keyness analysis: Nature, metrics and techniques. In C. Taylor & A. Marchi (Eds.), *Corpus Approaches to Discourse: A Critical Review* (pp. 225–258). Abingdon: Routledge.
- Gabrielatos, C., & Baker, P. (2008). Fleeing, sneaking, flooding: A corpus analysis of discursive constructions of refugees and asylum seekers in the UK press, 1996–2005. *Journal of English Linguistics*, 36(1), 5–38.
- Gabrielatos, C., McEnery, T., Diggle, P. J., & Baker, P. (2012). The peaks and troughs of corpus-based contextual analysis. *International Journal of Corpus Linguistics*, 17(2), 151–175.
- Gagolewski, M., Tartanus, B., IBM, Unicode Inc., & other contributors. (2017). *stringi: Character String Processing Facilities (Version 1.1.5)* [Computer software]. Available from <https://CRAN.R-project.org/package=stringi>
- GALE. (2018). The Times Digital Archive 1785–2012. Retrieved November 2018, from <https://www.gale.com/intl/c/the-times-digital-archive>
- Gale Cengage. (n.d.). The Times Digital Archive—Term Frequency. Retrieved August 2019, from http://gdc.galegroup.com/gdc/artemis/nGramViewer?p=TTDA&u=bham_uk#/1785/2013/DVI-Newspapers_ct/on//surveillance/
- Galtung, J., & Ruge, M. H. (1965). The structure of foreign news: The presentation of the Congo, Cuba and Cyprus crises in four Norwegian newspapers. *Journal of Peace Research*, 2(1), 64–90.

- Garside, R. (1987). The CLAWS word-tagging system. In R. Garside, G. Leech, & G. Sampson (Eds.), *The Computational Analysis of English: A Corpus-based Approach* (pp. 30–41). London: Longman.
- Gatto, M. (2014). *The Web As Corpus: Theory and Practice*. London: Bloomsbury.
- Godding, N. (2017, November 13). Hammerson's new Style app is snap happy. Retrieved July 2018, from <https://retailwatch.info/news/Hammerson%e2%80%99s-new-Style-app-is-snap-happy/8cf2b22b-8899-4233-974f-bbe2a376549e>
- Goffman, E. (1983). The Interaction Order: American Sociological Association, 1982 Presidential Address. *American Sociological Review*, 48(1), 1–17.
- Gordon, E., & de Souza e Silva, A. (2011). *Net Locality: Why Location Matters in a Networked World*. Chichester: Wiley-Blackwell.
- Görlach, M. (2002). A linguistic history of advertising, 1700–1890. In T. Fanego, B. Méndez-Naya, & E. Seoane (Eds.), *Sounds, Words, Texts and Change: Selected papers from 11 ICEHL, Santiago de Compostela, 7–11 September 2000. Volume 2* (pp. 83–104). Amsterdam: John Benjamins.
- Greenwald, G. (2015). *No Place to Hide: Edward Snowden, the NSA, and the U.S. Surveillance State*. Harmondsworth: Penguin. (Original work published 2014)
- Gregory, I. N., & Hardie, A. (2011). Visual GISing: Bringing together corpus linguistics and Geographical Information Systems. *Literary and Linguistic Computing*, 26(3), 297–314.
- Gries, S. Th. (2013). 50-something years of work on collocations: What is or should be next ... *International Journal of Corpus Linguistics*, 18(1), 137–166.
- Gries, S. Th., & Hilpert, M. (2008). The identification of stages in diachronic data: Variability-based neighbour clustering. *Corpora*, 3(1), 59–81.
- Grieve, J., Biber, D., Friginal, E., & Nekrasova, T. (2010). Variation among blogs: A multi-dimensional analysis. In A. Mehler, S. Sharoff, & M. Santini (Eds.), *Genres on the Web* (pp. 303–322). Dordrecht: Springer.
- Grieve, J., Nini, A., & Guo, D. (2017). Analyzing lexical emergence in Modern American English online. *English Language & Linguistics*, 21(1), 99–127.
- Grundmann, R., & Scott, M. (2014). Disputed climate science in the media: Do countries matter? *Public Understanding of Science*, 23(2), 220–235.
- Gupta, K. (2015). *Representation of the British Suffrage Movement*. London: Bloomsbury.
- Haggerty, K. D., & Ericson, R. V. (2000). The surveillant assemblage. *The British Journal of Sociology*, 51(4), 605–622.
- Hardie, A., & McEnery, T. (2010). On two traditions in corpus linguistics, and what they have in common. *International Journal of Corpus Linguistics*, 15(3), 384–394.
- Harding, L. (2014). *The Snowden Files: The Inside Story of the World's Most Wanted Man*. London: Guardian.
- Hennessey, A., Wiegand, V., Mahlberg, M., Tench, C. R., & Lentin, J. (2017). *CorporaCoCo: Corpora Co-Occurrence Comparison* (Version 1.1-0) [Computer software]. Available from <https://CRAN.R-project.org/package=CorporaCoCo>
- Herring, S. C. (2004). Slouching toward the ordinary: Current trends in Computer-Mediated Communication. *New Media & Society*, 6(1), 26–36.
- Herring, S. C., & Androutsopoulos, J. (2015). Computer-Mediated Discourse 2.0. In D. Tannen, H. E. Hamilton, & D. Schiffrin (Eds.), *The Handbook of Discourse Analysis* (2nd ed., pp. 127–151). Chichester: Wiley-Blackwell.
- Higgs, E. (2004). *The Information State in England: The Central Collection of Information on Citizens Since 1500*. Basingstoke: Palgrave Macmillan.

- Hoey, M. (2005). *Lexical Priming: A New Theory of Words and Language*. London: Routledge.
- Hoffmann, C. R. (2012). *Cohesive Profiling: Meaning and Interaction in Personal Weblogs*. Amsterdam: John Benjamins.
- Hood, S., & Forey, G. (2005). Introducing a conference paper: Getting interpersonal with your audience. *Journal of English for Academic Purposes*, 4(4), 291–306.
- Hornung, G., & Engemann, C. (2016). Einleitung [Introduction]. In C. Engemann & G. Hornung (Eds.), *Der digitale Bürger und seine Identität* [The digital citizen and his identity] (pp. 11–21). Baden-Baden: Nomos.
- Hunston, S. (2011). *Corpus Approaches to Evaluation: Phraseology and Evaluative Language*. London: Routledge.
- Huntley, S. J., Mahlberg, M., Wiegand, V., van Gennip, Y., Yang, H., Dean, R. S., & Brennan, M. L. (2018). Analysing the opinions of UK veterinarians on practice-based research using corpus linguistic and mathematical methods. *Preventive Veterinary Medicine*, 150, 60–69.
- Hyland, K. (2009). *Academic Discourse: English in a Global Context*. London: Continuum.
- Jaworski, A., & Thurlow, C. (2010). Introducing semiotic landscapes. In A. Jaworski & C. Thurlow (Eds.), *Semiotic Landscapes: Language, Image, Space* (pp. 1–40). London: Continuum.
- Jeffries, L., & Walker, B. (2018). *Keywords in the Press: The New Labour Years*. London: Bloomsbury.
- Jermyn, D. (2007). *Crime Watching: Investigating Real Crime TV*. London: I.B. Tauris.
- Jones, R. H. (2012). *Discourse Analysis: A Resource Book for Students*. London: Routledge.
- Jones, R. H. (2015). Surveillance. In A. Georgakopoulou & T. Spilioti (Eds.), *The Routledge Handbook of Language and Digital Communication* (pp. 408–411). Abingdon: Routledge.
- Jones, R. H. (2017). Surveillant landscapes. *Linguistic Landscape*, 3(2), 150–187.
- Jones, R. H., Chik, A., & Hafner, C. A. (2015). Introduction: Discourse analysis and digital practices. In R. H. Jones, A. Chik, & C. A. Hafner (Eds.), *Discourse and Digital Practices: Doing Discourse Analysis in the Digital Age* (pp. 1–17). New York, NY: Routledge.
- Jucker, A. H., & Berger, M. (2014). The development of discourse presentation in *The Times*, 1833–1988. *Media History*, 20(1), 67–87.
- Jurish, B. (2015). *DiaCollo: On the trail of diachronic collocations*. Presented at the CLARIN Annual Conference 2015, Wrocław. Retrieved July 2019 from <http://www.deutschestextarchiv.de/files/jurish2015diacollo-clarin.pdf>
- Jurish, B. (2018). Diachronic collocations, genre, and DiaCollo. In R. J. Whitt (Ed.), *Diachronic Corpora, Genre, and Language Change*. Amsterdam: John Benjamins.
- Jurish, B., Geyken, A., & Werneke, T. (2016, March 7). *DiaCollo: Diachronen Kollokationen auf der Spur* [*DiaCollo: Tracking diachronic collocations*]. In *Konferenzabstracts DHd 2016: Modellierung – Vernetzung – Visualisierung* [Modelling – Networking – Visualising] (pp. 172–175). Leipzig: Universität Leipzig. Retrieved August 2019, from <http://dhd2016.de/boa.pdf>
- Kehoe, A., & Gee, M. (2009). Weaving web data into a diachronic corpus patchwork. In A. Renouf & A. Kehoe (Eds.), *Corpus Linguistics: Refinements and Reassessments* (pp. 255–279). Amsterdam: Rodopi.
- Khan, K. (2019). *Becoming a Citizen: Linguistic Trials and Negotiations in the UK*. Bloomsbury.

- Kilgariff, A. (2005). Language is never, ever, ever, random. *Corpus Linguistics and Linguistic Theory*, 1(2), 263–276.
- King, P. (2004). *Private Dwelling: Contemplating the Use of Housing*. London: Routledge.
- Kleinke, S. (2012). Foundations of pragmatics. *Anglia - Zeitschrift Für Englische Philologie*, 130(3), 420–426.
- Knight, D., Adolphs, S., Tennent, P., & Carter, R. (2008). The Nottingham Multi-Modal Corpus: A demonstration. In *Proceedings of the 6th Language Resources and Evaluation Conference (LREC), Palais Des Congr s, Marrakech, Morocco, 28-30th May 2008* (pp. 1–7). Retrieved July, 2019 from <https://orca.cf.ac.uk/78855/>
- Koskela, H. (2000). ‘The gaze without eyes’: Video-surveillance and the changing nature of urban space. *Progress in Human Geography*, 24(2), 243–265.
- Koteyko, N. (2006). Corpus Linguistics and the Study of Meaning in Discourse. *The Linguistics Journal*, 1(2), 132–157.
- Koteyko, N. (2007). A diachronic approach to meaning: English loanwords in Russian opposition discourse. *Corpora*, 2(1), 65–95.
- Krishnamurthy, R. (2018). Modes of analysis: An autoergography of corpus linguistics from lexis to discourse. In A.  erm kov  & M. Mahlberg (Eds.), *The Corpus Linguistics Discourse: In Honour of Wolfgang Teubert* (pp. 223–253). Amsterdam: John Benjamins.
- Kroener, I. (2013). “Caught on Camera”: The media representation of video surveillance in relation to the 2005 London Underground bombings. *Surveillance & Society*, 11(1/2), 121–133.
- Landry, R., & Bourhis, R. Y. (1997). Linguistic landscape and ethnolinguistic vitality: An empirical study. *Journal of Language and Social Psychology*, 16(1), 23–49.
- Langacker, R. W. (2008). *Cognitive Grammar: A Basic Introduction*. Oxford: Oxford University Press.
- Lee, D. (2002). Genres, registers, text types, domains and styles: Clarifying the concepts and navigating a path through the BNC jungle. *Language and Computers*, 42(1), 247–292.
- Leech, G. (1985). Stylistics. In T. A. van Dijk (Ed.), *Discourse and Literature: New Approaches to the Analysis of Literary Genres* (pp. 39–57). Amsterdam: John Benjamins.
- Leech, G., & Short, M. (2007). *Style in Fiction: A Linguistic Introduction to English Fictional Prose* (2nd ed.). Harlow: Pearson.
- Leurdijk, A., Nieuwenhuis, O., & Poel, M. (2014). The newspaper industry. In G. De Prato, E. Sanz, & J. P. Simon (Eds.), *Digital Media Worlds: The New Economy of Media* (pp. 147–162). Houndmills: Palgrave Macmillan.
- LexisNexis. (2019). Nexis : News Search. Retrieved May 2019, from https://www.nexis.com/search/editSearch.do?formBeanKey=37_T28746388036&BC T=G0
- Lou, J. J. (2014). Locating the power of place in space: A geosemiotic approach to context. In J. Flowerdew (Ed.), *Discourse in Context: Contemporary Applied Linguistics* (pp. 205–224). London: Bloomsbury.
- LSE Identity Project. (2005). *The Identity Project: An assessment of the UK Identity Cards Bill and its Implications*. London: LSE Department of Information Systems. Retrieved August 2019, from <http://www.lse.ac.uk/management/research/identityproject/identityreport.pdf>.
- Lupton, D. (2015). *Digital Sociology*. Abingdon: Routledge.

- Lyon, D. (2001). *Surveillance Society: Monitoring Everyday Life*. Buckingham: Open University Press.
- Lyon, D. (2007). *Surveillance Studies: An Overview*. Cambridge: Polity.
- Lyon, D. (2009). *Identifying Citizens: ID Cards as Surveillance*. Cambridge: Polity.
- Lyon, D. (2010). Liquid surveillance: The contribution of Zygmunt Bauman to surveillance studies. *International Political Sociology*, 4(4), 325–338.
- Lyon, D. (2014). Surveillance, Snowden, and Big Data: Capacities, consequences, critique. *Big Data & Society*, 1(2), 1–13.
- Lyons, K. (2018). *From Street to Screen: Linguistic Productions of Place in San Francisco's Mission District* (Unpublished doctoral dissertation). University of Illinois at Urbana-Champaign, Champaign, IL. Retrieved July 2019, from <http://hdl.handle.net/2142/100965>
- MacDonald, M., & Hunter, D. (2013a). Security, population and governmentality: UK counter-terrorism discourse (2007-2011). *Critical Approaches to Discourse Analysis Across Disciplines*, 7(1), 123–140.
- MacDonald, M., & Hunter, D. (2013b). The discourse of Olympic security: London 2012. *Discourse & Society*, 24(1), 66–88.
- MacDonald, M., Hunter, D., & O'Regan, J. (2013). Citizenship, community, and counter-terrorism: UK security discourse, 2001-2011. *Journal of Language and Politics*, 12(3), 445–473.
- Mahlberg, M. (2005). *English General Nouns: A Corpus Theoretical Approach*. Amsterdam: John Benjamins.
- Mahlberg, M. (2007a). Clusters, key clusters and local textual functions in Dickens. *Corpora*, 2(1), 1–31.
- Mahlberg, M. (2007b). Corpus stylistics: Bridging the gap between linguistic and literary studies. In M. Hoey, M. Mahlberg, M. Stubbs, & W. Teubert (Eds.), *Text, Discourse and Corpora: Theory and Analysis*. London: Continuum.
- Mahlberg, M. (2007c). Lexical items in discourse: Identifying local textual functions of sustainable development. In M. Hoey, M. Mahlberg, M. Stubbs, & W. Teubert (Eds.), *Text, Discourse and Corpora: Theory and Analysis* (pp. 191–218). London: Continuum.
- Mahlberg, M. (2013). *Corpus Stylistics and Dickens's Fiction*. London: Routledge.
- Mahlberg, M. (2014). Corpus linguistics and discourse analysis. In K. P. Schneider & A. Barron (Eds.), *Pragmatics of Discourse* (pp. 215–238). Berlin: De Gruyter Mouton.
- Mahlberg, M., & McIntyre, D. (2011). A case for corpus stylistics: Ian Fleming's *Casino Royale*. *English Text Construction*, 4(2), 204–227.
- Mahlberg, M., & O'Donnell, M. B. (2008). *A fresh view of the structure of hard news stories*. Presented at the 19th European Systemic Functional Linguistics Conference and Workshop, Saarbrücken. Retrieved August 2019, from https://publikationen.sulb.uni-saarland.de/bitstream/20.500.11880/23572/1/Mahlberg_ODonnell_form.pdf
- Mahlberg, M., & Wiegand, V. (2018). Corpus stylistics, norms and comparisons: Studying speech in *Great Expectations*. In R. Page, B. Busse, & N. Nørgaard (Eds.), *Rethinking Language, Text and Context: Interdisciplinary Research in Stylistics in Honour of Michael Toolan* (pp. 123–143). London: Routledge.
- Mair, C. (2006). *Twentieth-Century English: History, Variation, and Standardization*. Cambridge: Cambridge University Press.
- Malinowski, D. (2009). Authorship in the linguistic landscape: A multimodal-performative view. In D. Gorter & E. Shohamy (Eds.), *Linguistic Landscape: Expanding the Scenery* (pp. 107–125). New York: Routledge.

- Marchi, A. (2018). Dividing up the data: Epistemological, methodological and practical impact of diachronic segmentation. In C. Taylor & A. Marchi (Eds.), *Corpus Approaches to Discourse: A Critical Review* (pp. 174–196). Abingdon, Oxon: Routledge.
- Marchi, A., & Taylor, C. (2018). Introduction: Partiality and reflexivity. In C. Taylor & A. Marchi (Eds.), *Corpus Approaches to Discourse: A Critical Review* (pp. 1–15). Abingdon: Routledge.
- Mastropierro, L. (2017). *Corpus Stylistics in Heart of Darkness and its Italian Translations*. London: Bloomsbury.
- McEnery, T. (2006). *Swearing in English: Bad Language, Purity and Power from 1586 to the Present*. London: Routledge.
- McEnery, T., & Baker, P. (Eds.). (2015). *Corpora and Discourse Studies*. Palgrave Macmillan.
- McEnery, T., & Hardie, A. (2012). *Corpus Linguistics: Method, Theory and Practice*. Cambridge: Cambridge University Press.
- McEnery, T., McGlashan, M., & Love, R. (2015). Press and social media reaction to ideologically inspired murder: The case of Lee Rigby. *Discourse & Communication*, 9(2), 237–259.
- McEnery, T., & Xiao, R. (2010). *Corpus-Based Contrastive Studies of English and Chinese*. London: Routledge.
- Mehrabov, I. (2015). Exploring terra incognita: Mapping surveillance studies from the perspective of media and communication research. *Surveillance & Society*, 13(1), 117–126.
- Meinel, C., Berger, P., Bross, J., & Hennig, P. (2015). *Blogosphere and its Exploration*. Heidelberg: Springer.
- Mills, S. (2004). *Discourse: The New Critical Idiom* (2nd ed.). London: Routledge.
- Molino, A. (2019). Corporate identity and its variation over time: A corpus-assisted study of self-presentation strategies in Vodafone's Sustainability Reports. In V. Wiegand & M. Mahlberg (Eds.), *Corpus Linguistics, Context and Culture* (pp. 75–107). Berlin: De Gruyter Mouton.
- Murakami, A., Thompson, P., Hunston, S., & Vajn, D. (2017). 'What is this corpus about?' Using topic modeling to explore a specialized corpus. *Corpora*, 12(2), 243–277.
- Myers, G. (2006). 'Where are you from?': Identifying place. *Journal of Sociolinguistics*, 10(3), 320–343.
- Myers, G. (2010). *Discourse of Blogs and Wikis*. London: Continuum.
- Nolte, I., Ancarno, C., & Jones, R. (2018). Using corpus methods to explore a survey on inter-religious relations in Yorubaland, southwest Nigeria. *Corpora*, 13(1), 27–64.
- Norris, C., McCahill, M., & Wood, D. (2004). The growth of CCTV: A global perspective on the international diffusion of video surveillance in publicly accessible space. *Surveillance & Society*, 2(2/3), 110–135.
- O'Donnell, M. B., Scott, M., Mahlberg, M., & Hoey, M. (2012). Exploring text-initial words, clusters and concgrams in a newspaper corpus. *Corpus Linguistics and Linguistic Theory*, 8(1), 73–101.
- O'Keeffe, A. (2012). Corpora and media studies. In K. Hyland, M. H. Chau, & M. Handford (Eds.), *Corpus Applications in Applied Linguistics*. London: Continuum.
- Partington, A., Duguid, A., & Taylor, C. (2013). *Patterns and Meanings in Discourse: Theory and Practice in Corpus-Assisted Discourse Studies (CADS)*. Amsterdam: John Benjamins.
- Pearson, J. (1998). *Terms in Context*. Amsterdam: John Benjamins.

- Pecorari, D. (2009). Formulaic language in biology: A topic-specific investigation. In M. Charles, D. Pecorari, & S. Hunston (Eds.), *Academic Writing: At the Interface of Corpus and Discourse* (pp. 91–104). London: Continuum.
- Phillips, M. A. (1985). *Aspects of Text Structure: An Investigation of the Lexical Organisation of Text*. Amsterdam: North-Holland.
- Potts, A., Bednarek, M., & Caple, H. (2015). How can computer-based methods help researchers to investigate news values in large datasets? A corpus linguistic study of the construction of newsworthiness in the reporting on Hurricane Katrina. *Discourse & Communication*.
- Prentice, S., Rayson, P., & Taylor, P. J. (2012). The language of Islamic extremism: Towards an automated identification of beliefs, motivations and justifications. *International Journal of Corpus Linguistics*, 17(2), 259–286.
- Puschmann, C. (2010). *The Corporate Blog as an Emerging Genre of Computer-mediated Communication: Features, Constraints, Discourse Situation*. Göttingen: Universitätsverlag Göttingen.
- R Core Team. (2016). *R: A Language and Environment for Statistical Computing* [Computer software]. Available from <http://www.R-project.org/>
- Rayson, P. (2003). *Matrix: A Statistical Method and Software Tool for Linguistic Analysis Through Corpus Comparison* (Unpublished doctoral dissertation). University of Lancaster, Lancaster. Retrieved September 2017, from <http://eprints.lancs.ac.uk/12287/>
- Rayson, P. (2008). From key words to key semantic domains. *International Journal of Corpus Linguistics*, 13(4), 519–549.
- Rayson, P. (2009). *Wmatrix: A Web-based Corpus Processing Environment*. [Computer software]. Available from <http://ucrel.lancs.ac.uk/wmatrix/>
- Rayson, P. (n.d.). *Log-likelihood and effect size calculator*. Available from <http://ucrel.lancs.ac.uk/llwizard.html>
- Rayson, P., Archer, D., Piao, S., & McEnery, T. (2004). The UCREL semantic analysis system. In *Proceedings of the Workshop on Beyond Named Entity Recognition Semantic Labelling for NLP Tasks in Association with LREC 2004*. Lisbon. Retrieved July 2019, from https://www.lancaster.ac.uk/staff/rayson/publications/usas_lrec04ws.pdf
- Research and Development Unit for English Studies. (2017). *WebCorp: The Web as Corpus* [Online corpus tool]. Available from <http://www.webcorp.org.uk/live/index.jsp>
- Rettberg, J. W. (2008). *Blogging*. Cambridge: Polity Press.
- RStudio Team. (2015). *RStudio: Integrated Development for R* [Computer software]. Available from <http://www.rstudio.com/>
- Schmid, H. (2008). Tokenizing and part-of-speech tagging. In A. Lüdeling & M. Kytö (Eds.), *Corpus Linguistics: An International Handbook* (Vol. 1, pp. 527–551). Berlin: Mouton de Gruyter.
- Schröter, M., & Storjohann, P. (2015). Patterns of discourse semantics. *Pragmatics and Society*, 6(1), 43–66.
- Schulze, M. (2015). Patterns of surveillance legitimization: The German discourse on the NSA scandal. *Surveillance & Society*, 13(2), 197–217.
- Scollon, R. (2001). *Mediated Discourse: The Nexus of Practice*. London: Routledge.
- Scollon, R., & Scollon, S. W. (2003). *Discourses in Place: Language in the Material World*. London: Routledge.
- Scollon, R., & Scollon, S. W. (2004). *Nexus Analysis: Discourse and the Emerging Internet*. London: Routledge.
- Scott, M. (1997). PC analysis of key words—And key key words. *System*, 25(2), 233–245.

- Scott, M. (2016). *WordSmith Tools* (Version 7) [Computer software]. Stroud: Lexical Analysis Software.
- Scott, M. (n.d.). Downloading BNC or ANC word lists. Retrieved July 2019, from <https://lexically.net/wordsmith/index.html>
- Scott, M., & Tribble, C. (2006). *Textual Patterns: Key Words and Corpus Analysis in Language Education*. Amsterdam: John Benjamins.
- Seale, C., Boden, S., Williams, S., Lowe, P., & Steinberg, D. (2007). Media constructions of sleep and sleep disorders: A study of UK national newspapers. *Social Science & Medicine*, 65(3), 418–430.
- Seale, C., Ziebland, S., & Charteris-Black, J. (2006). Gender, cancer experience and internet use: A comparative keyword analysis of interviews and online cancer support groups. *Social Science & Medicine*, 62(10), 2577–2590.
- Selinker, L., Trimble, R. M. T., & Trimble, L. (1976). On Reading English for Science and Technology: Presuppositional Rhetorical Information in the Discourse. In J. C. Richards (Ed.), *Teaching English for science and technology: Selected papers from the RELC Seminar on the Teaching and Learning of English for Scientific and Technological Purposes in Southeast Asia, Singapore, 21-25 April 1975* (pp. 37–67). Singapore: Singapore University Press for Seameo Regional English Language Centre.
- Sinclair, J. (Ed.). (1987). *Looking up: An Account of the COBUILD Project in Lexical Computing and the Development of the Collins COBUILD English Language Dictionary*. London: Collins.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Sinclair, J. (2004a). Intuition and annotation – the discussion continues. In K. Aijmer & B. Altenberg (Eds.), *Advances in Corpus Linguistics: Papers from the 23rd International Conference on English Language Research on Computerized Corpora (ICAME 23), Göteborg 22-26 May 2002* (pp. 39–59). Amsterdam: Rodopi.
- Sinclair, J. (2004b). *Trust the Text: Language, Corpus and Discourse*. London: Routledge.
- Sinclair, J. (2005). Corpus and text: Basic principles. In M. Wynne (Ed.), *Developing Linguistic Corpora: A Guide to Good Practice*. Retrieved August 2019, from <https://ota.ox.ac.uk/documents/creating/dlc/chapter1.htm>
- Sinclair, J., Jones, S., & Daley, R. (2004). *English Collocation Studies: The OSTI Report* (R. Krishnamurthy, Ed.). London: Bloomsbury. (Original work published 1970)
- Sixty9°. (2017, October 24). Innovative new style seeker tool launches in Leicester's Highcross with living mannequins in centre. Retrieved July 25, 2018, from <http://69-degrees.co.uk/leicester/innovative-new-style-seeker-tool-launches-highcross-leicester-become-shopping-saviour/>
- Slater, D. (1998). Public/private. In C. Jenks (Ed.), *Core Sociological Dichotomies* (pp. 138–150). London: Sage.
- Solid Documents Limited. (2015). *Solid Converter Mac (Trial Version)* (Version 2) [Computer software]. Available from <http://www.mac-pdf-converter.com>
- Stubbs, M. (2001). *Words and Phrases: Corpus Studies of Lexical Semantics*. Oxford: Blackwell.
- Swales, J. (1984). *Episodes In ESP: A Source and Reference Book on the Development of English for Science and Technology*. Oxford: Pergamon.
- Taylor, C. (2008). What is corpus linguistics? What the data says. *ICAME Journal*, 32, 179–200.
- Taylor, C., & Marchi, A. (Eds.). (2018). *Corpus Approaches to Discourse: A Critical Review*. Abingdon: Routledge.

- Teich, E., & Fankhauser, P. (2009). Exploring a corpus of scientific texts using data mining. *Language and Computers*, 71(1), 233–247.
- Teich, E., & Holtz, M. (2009). Scientific registers in contact: An exploration of the lexicogrammatical properties of interdisciplinary discourses. *International Journal of Corpus Linguistics*, 14(4), 524–548.
- Teubert, W. (2001). Corpus linguistics and lexicography. *International Journal of Corpus Linguistics*, 6(3), 125–153.
- Teubert, W. (2005a). Evaluation and its discontents. In E. Tognini-Bonelli & G. Del Lungo Camiciotti (Eds.), *Strategies in Academic Discourse* (pp. 185–204). Amsterdam: John Benjamins.
- Teubert, W. (2005b). My version of corpus linguistics. *International Journal of Corpus Linguistics*, 10(1), 1–13.
- Teubert, W. (2007a). Natural and human rights, work and property in the discourse of Catholic social doctrine. In M. Hoey, M. Mahlberg, M. Stubbs, & W. Teubert (Eds.), *Text, Discourse and Corpora: Theory and Analysis* (pp. 89–126). London: Continuum.
- Teubert, W. (2007b). Parole-linguistics and the diachronic dimension of discourse. In M. Hoey, M. Mahlberg, M. Stubbs, & W. Teubert (Eds.), *Text, Discourse and Corpora: Theory and Analysis* (pp. 57–88). London: Continuum.
- Teubert, W. (2010). *Meaning, Discourse and Society*. Cambridge: Cambridge University Press.
- Teubert, W. (2012). Democracy and web-based dialogue. In F. Cooren & A. Létourneau (Eds.), *(Re)presentations and Dialogue* (pp. 99–123). Amsterdam: John Benjamins.
- Teubert, W. (2019). Corpus linguistics: Widening the remit. In V. Wiegand & M. Mahlberg (Eds.), *Corpus Linguistics, Context and Culture* (pp. 137–161). Berlin: De Gruyter.
- Teubert, W., & Čermáková, A. (2004). Directions in corpus linguistics. In M. A. K. Halliday, W. Teubert, C. Yallop, & A. Čermáková (Eds.), *Lexicology and Corpus Linguistics* (pp. 113–166). London: Continuum.
- The Editors. (2009, January 30). Vol 6 No 1 (2009): Relaunch Issue: Revisiting Video Surveillance | Surveillance & Society. Retrieved July 2019, from <https://ojs.library.queensu.ca/index.php/surveillance-and-society/issue/view/Relaunch>
- The International Football Association Board. (2018, March 3). Historic step for greater fairness in football. Retrieved January 2019, from <http://theifab.com>
- The Phoenix Newspaper. (2017, November 2). Bullring help to Style Birmingham with StyleSeeker Tool. Retrieved July 2018, from <https://thephoenixnewspaper.com/bullring-help-style-birmingham-styleseeker-tool>
- Thurlow, C., & Mroczek, K. (Eds.). (2011a). *Digital Discourse: Language in the New Media*. Oxford: Oxford University Press.
- Thurlow, C., & Mroczek, K. (2011b). Introduction. In C. Thurlow & K. Mroczek (Eds.), *Digital Discourse: Language in the New Media*. Oxford: Oxford University Press.
- Thurman, N., & Fletcher, R. (2018). Are newspapers heading toward post-print obscurity? *Digital Journalism*, 6(8), 1003–1017.
- Titak, A., & Roberson, A. (2013). Dimensions of web registers: An exploratory multi-dimensional comparison. *Corpora*, 8(2), 235–260.
- Tognini-Bonelli, E. (2001). *Corpus Linguistics at Work*. Amsterdam: John Benjamins.
- Topshop. (n.d.). Privacy & Cookies. Retrieved July 2018, from <http://www.topshop.com/en/tsuk/category/privacy-cookies-7725029/home>
- Triki, N. (2019). Revisiting the metadiscursive aspect of definitions in academic writing. *Journal of English for Academic Purposes*, 37, 104–116.

- Turnbull, J. (2013). Building, enhancing and defending reputation in a corporate website. In R. Salvi & J. Bowker (Eds.), *Space, Time and the Construction of Identity: Discursive Indexicality in Cultural, Institutional and Professional Fields* (pp. 293–317). Bern: Peter Lang.
- Turow, J. (2011). *The Daily You: How the New Advertising Industry is Defining your Identity and your Worth*. New Haven: Yale University Press.
- Virtanen, T. (2009). Discourse linguistics meets corpus linguistics: Theoretical and methodological issues in the troubled relationship. In A. Renouf & A. Kehoe (Eds.), *Corpus Linguistics: Refinements and Reassessments* (pp. 49–65). Amsterdam: Rodopi.
- Walby, K. (2005). How closed-circuit television surveillance organizes the social: An institutional ethnography. *The Canadian Journal of Sociology / Cahiers Canadiens de Sociologie*, 30(2), 189–214.
- Webster, W. (2009). CCTV policy in the UK: Reconsidering the evidence base. *Surveillance & Society*, 6(1), 10–22.
- Weingart, P., Engels, A., & Pansegrau, P. (2000). Risks of communication: Discourses on climate change in science, politics, and the mass media. *Public Understanding of Science*, 9(3), 261–283.
- Weisser, M. (2016). *Practical Corpus Linguistics: An Introduction to Corpus-based Language Analysis*. Oxford: Wiley-Blackwell.
- West Midlands Growth Company. (n.d.). Shopping centres. Retrieved July 2018, from <http://visitbirmingham.com/what-to-do/shopping/shopping-centres/>
- Whiteley, S., & Canning, P. (2017). Reader response research in stylistics. *Language and Literature*, 26(2), 71–87.
- Whitley, E. A. (2011). Perceptions of government technology, surveillance and privacy: The UK Identity Cards Scheme. In B. J. Goold & D. Neyland (Eds.), *New Directions in Surveillance and Privacy* (pp. 133–156). Abingdon: Routledge. (Original work published 2009)
- Whitley, E. A., Hosein, I. R., Angell, I. O., & Davies, S. (2007). Reflections on the academic policy analysis process and the UK Identity Cards Scheme. *The Information Society*, 23(1), 51–58.
- Wiegand, V. (2019, May). *Surveillance and duty: A study of newspaper advertisements from 1800–1945*. Presented at the CHINED VII, Barcelona.
- Wiegand, V., Hennessey, A., Tench, C. R., & Mahlberg, M. (2017a, July 24). *A cookbook of co-occurrence comparison techniques and how they relate to the subtleties in your research question*. Presented at the 9th International Corpus Linguistics Conference, Birmingham. Retrieved July 2019, from <https://www.birmingham.ac.uk/Documents/college-artslaw/corpus/conference-archives/2017/general/paper362.pdf>
- Wiegand, V., Hennessey, A., Tench, C. R., & Mahlberg, M. (2017b, May 24). *Comparing co-occurrences between corpora*. Presented at the 38th ICAME conference, Prague.
- Wiegand, V., & Mahlberg, M. (2019). Introduction. In V. Wiegand & M. Mahlberg (Eds.), *Corpus Linguistics, Context and Culture* (pp. 1–9). Berlin: De Gruyter Mouton.
- Williams, G. (1998). Collocational networks: Interlocking patterns of lexis in a corpus of plant biology research articles. *International Journal of Corpus Linguistics*, 3(1), 151–171.
- Williams, G. (2002). In search of representativity in specialised corpora: Categorisation through collocation. *International Journal of Corpus Linguistics*, 7(1), 43–64.
- Williams, Raymond. (2011). *Keywords: A Vocabulary of Culture and Society* (2nd ed.). Abingdon: Routledge. (Original work published 1976)

- Winseck, D. (2003). Netscapes of power: Convergence, network design, walled gardens, and other strategies of control in the information age. In D. Lyon (Ed.), *Surveillance as Social Sorting: Privacy, Risk, and Digital Discrimination* (pp. 176–198). London: Routledge.
- Wynne, M. (2008). Searching and concordancing. In A. Lüdeling & M. Kytö (Eds.), *Corpus Linguistics: An International Handbook* (Vol. 1, pp. 706–737). Berlin: Mouton de Gruyter.
- Zanchetta, E., Baroni, M., & Bernardini, S. (2011, July). *Corpora for the masses: The BootCaT front-end*. Presented at the Corpus Linguistics 2011 Conference, Birmingham. Retrieved July 2019, from <https://www.birmingham.ac.uk/documents/college-artslaw/corpus/conference-archives/2011/abs-55.pdf>
- Zanchetta, E., & Gaspari, F. (2018, February 7). BootCaT front-end tutorial – Part 2. Retrieved March 2019, from http://docs.sslmit.unibo.it/doku.php?id=bootcat:tutorials:basic_2
- Zinn, J. O., & McDonald, D. (2017). *Risk in The New York Times (1987–2014): A Corpus-based Exploration of Sociological Theories*. Cham: Springer.
- Zou, H., & Hyland, K. (2019). Reworking research: Interactions in academic articles and blogs. *Discourse Studies*. Advance online publication. Retrieved July 2019, from <https://doi.org/10.1177/1461445619866983>
- Zurawski, N. (Ed.). (2007). Einleitung [Introduction]. In N. Zurawski (Ed.), *Surveillance Studies: Perspektiven eines Forschungsfeldes* [Surveillance Studies: Perspectives of a research field] (pp. 7–24). Opladen: Barbara Budrich.

References of cited examples from the S&S Corpus

- Abu-Laban, Y. (2015). Gendering surveillance studies: The empirical and normative promise of feminist methodology. *Surveillance & Society*, 13(1), 44–56.
- Adey, P. (2003). Secured and sorted mobilities: Examples from the airport. *Surveillance & Society*, 1(4), 500–519.
- Bauer, S., & Olsén, J. E. (2009). Observing the others, watching over oneself: Themes of medical surveillance in post-panoptic society. *Surveillance & Society*, 6(2), 116–127.
- Bennett, C. J., & Regan, P. M. (2003). Surveillance and mobilities. *Surveillance & Society*, 1(4), 449–455.
- Boersma, K. (2013). Liminal Surveillance. Intensified use of an existing CCTV system during a local event. *Surveillance & Society*, 11(1/2), 106–120.
- Braverman, I. (2012). Zooveillance: Foucault goes to the zoo. *Surveillance & Society*, 10(2), 119–133.
- Collister, L. B. (2014). Surveillance and community: Language policing and empowerment in a World of Warcraft Guild. *Surveillance & Society*, 12(3), 337–348.
- Corones, A., & Hardy, S. (2009). En-gendered surveillance: Women on the edge of a watched cervix. *Surveillance & Society*, 6(4), 388–397.
- Edmond, G., & Roque, M. S. (2013). Justicia's Gaze: Surveillance, evidence and the criminal trial. *Surveillance & Society*, 11(3), 252–271.
- Fotel, T., & Thomsen, T. U. (2003). The surveillance of children's mobility. *Surveillance & Society*, 1(4), 535–554.
- French, M. A. (2009). Woven of war-time fabrics: The globalization of public health surveillance. *Surveillance & Society*, 6(2), 101–115.
- Germain, S., Dumoulin, L., & Douillet, A.-C. (2013). A prosperous 'business'. The success of CCTV through the eyes of international literature. *Surveillance & Society*, 11(1/2), 134–147.
- Gray, M. (2003). Urban surveillance and panopticism: Will we recognize the facial recognition society? *Surveillance & Society*, 1(3), 314–330.
- Green, N., & Zurawski, N. (2015). Surveillance and ethnography: Researching surveillance as everyday life. *Surveillance & Society*, 13(1), 27–43.
- Hier, S. P. (2003). Probing the Surveillant Assemblage: On the dialectics of surveillance practices as processes of social control. *Surveillance & Society*, 1(3), 399–411.
- Introna, L. D. (2003). Workplace surveillance 'is' unethical and unfair. *Surveillance & Society*, 1(2), 210–216.
- Keiber, J. (2015). Surveillance hegemony. *Surveillance & Society*, 13(2), 168–181.
- Klauser, F. R. (2004). A comparison of the impact of protective and preservative video surveillance on urban territoriality: The case of Switzerland. *Surveillance & Society*, 2(2/3), 145–160.
- Klauser, F. R. (2009). 'Lost' Surveillance Studies: A critical review of French work on CCTV. *Surveillance & Society*, 6(1), 23–31.
- Maki, K. (2011). Neoliberal deviants and surveillance: Welfare recipients under the watchful eye of Ontario Works. *Surveillance & Society*, 9(1/2), 47–63.
- Martin, A. K., Brakel, R. E. van, & Bernhard, D. J. (2009). Understanding resistance to digital surveillance: Towards a multi-disciplinary, multi-actor framework. *Surveillance & Society*, 6(3), 213–232.

- Marx, G. T. (2002). What's new about the "new surveillance"? Classifying for change and continuity. *Surveillance & Society*, 1(1), 9–29.
- McIntosh, I., Punch, S., Dorrer, N., & Emond, R. (2010). 'You don't have to be watched to make your toast': Surveillance and food practices within residential care. *Surveillance & Society*, 7(3/4), 290–303.
- Mehrabov, I. (2015). Exploring terra incognita: Mapping surveillance studies from the perspective of media and communication research. *Surveillance & Society*, 13(1), 117–126.
- Monahan, T., & Wall, T. (2007). Somatic surveillance: Corporeal control through information networks. *Surveillance & Society*, 4(3), 154–173.
- Murakami Wood, D. (2009a). An olive branch to Hier and Greenberg. *Surveillance & Society*, 7(1), 76–77.
- Murakami Wood, D. (2009b). Situating surveillance studies. Sean Hier and Josh Greenberg's *The Surveillance Studies Reader*, and David Lyon's *Surveillance Studies: An Overview*. *Surveillance & Society*, 6(1), 52–61.
- Murakami Wood, D., & Wright, S. (2015). Before and after Snowden. *Surveillance & Society*, 13(2), 132–138.
- Negishi, K. (2013). From surveillant text to surveilling device: The face in urban transit spaces. *Surveillance & Society*, 11(3), 324–333.
- Phillips, D. J. (2015). Work and play at the threshold of legibility: Theatre as method and pedagogy in surveillance research. *Surveillance & Society*, 13(1), 57–77.
- Saltes, N. (2013). 'Abnormal' bodies on the borders of inclusion: Biopolitics and the paradox of disability surveillance. *Surveillance & Society*, 11(1/2), 55–73.
- Stanton, J. M., & Stam, K. R. (2003). Information technology, privacy, and power within organizations: A view from boundary theory and social exchange perspectives. *Surveillance & Society*, 1(2), 152–190.
- Stoddart, E. (2014). Challenging 'Just Surveillance Theory': A response to Kevin Macnish's 'Just Surveillance? Towards a Normative Theory of Surveillance.' *Surveillance & Society*, 12(1), 158–163.
- Timan, T., & Oudshoorn, N. (2012). Mobile cameras as new technologies of surveillance? How citizens experience the use of mobile cameras in public nightscapes. *Surveillance & Society*, 10(2), 167–181.
- van der Velden, L. (2015). Leaky Apps and Data Shots: Technologies of Leakage and Insertion in NSA-Surveillance. *Surveillance & Society*, 13(2), 182–196.
- Webster, W. (2009). CCTV policy in the UK: Reconsidering the evidence base. *Surveillance & Society*, 6(1), 10–22.
- Wells, H., & Wills, D. (2009). Individualism and identity: Resistance to speed cameras in the UK. *Surveillance & Society*, 6(3), 259–274.
- Whelan, C. (2014). Surveillance, security and sporting mega events: Toward a research agenda on the organisation of security networks. *Surveillance & Society*, 11(4), 392–404.
- Wiebe, S. (2009). Producing Bodies and Borders: A review of immigrant medical examinations in Canada. *Surveillance & Society*, 6(2), 128–141.
- Williams, Robin, & Johnson, P. (2004). Circuits of surveillance? *Surveillance & Society*, 2(1), 1–14.

References of cited examples from the Surveillance Blog Corpus

- Amberhawk. (2016, January 5). Draft Investigatory Powers Bill ignores data protection when collecting bulk personal datasets. Retrieved August 2018, from <http://amberhawk.typepad.com/amberhawk/2016/01/draft-investigatory-powers-bill-ignores-data-protection-when-collecting-bulk-personal-datasets.html>
- Amberhawk. (n.d.). Retrieved March 30, 2019, from <https://amberhawk.typepad.com/amberhawk/>
- ASC. (2014, July 11). Old War Movies: THE ENIGMA OF DRONES. Retrieved July 2019, from <http://moviephilippines.blogspot.com/2014/07/the-enigma-of-drones.html>
- Bhadra, A. (2010, April 30). Arindam Bhadra: CCTV Index / Video Surveillance Guide. Retrieved July 2019, from <http://arindamcctvaccesscontrol.blogspot.com/2009/12/cctv-index.html>
- Big Gav. (2007, June 11). Peak Energy: The Shockwave Rider. Retrieved August 2018, from <http://peakenergy.blogspot.com/2007/06/shockwave-rider.html>
- BIGMART. (2005, July 28). BIGMART's Privacy solution site: Mid July - 2005 Newsletter. Retrieved August 2018, from <http://offshoreonline.blogspot.com/2005/07/mid-july-2005-newsletter.html>
- CFC. (2010, November 2). Airport Digital Strip Searches or Aggressive Pat Downs? Retrieved August 2018, from <http://consumercal.blogspot.com/2010/11/airport-digital-strip-searches-or.html>
- Coronare Modestus Faust. (2013, December 10). The "Nothing To Hide" Defense—A Not Altogether Thoughtful Position That Forfeits Freedoms In An NSA Age. Retrieved August 2018, from <https://spfaust.wordpress.com/2013/12/10/the-nothing-to-hide-defense-a-not-altogether-thoughtful-position-that-forfeits-freedoms-in-an-nsa-age/>
- Dwight, R. J. (2009, April 21). Catholic Prophecy Today: Catholic Prophecy Today. Retrieved April 2019, from http://markbeast.blogspot.com/2009/04/prophecy-and-current-events_21.html
- Guillaume-sam, J. (2007, August 21). Silence is Betrayal: THE METHOD. Retrieved June 2018, from <http://antineworldorder.blogspot.com/2007/08/method.html>
- Hissil. (2009, August 13). Hissil's Blogspot - Exposing Disinfo Shills and Trolls: The David Icke Guide To The Global Conspiracy 3/4. Retrieved August 2018, from http://hissil.blogspot.com/2009/08/david-icke-guide-to-global-conspiracy_13.html
- JQ. (2005, April 26). Response to "New Surveillance Cameras can detect behaviors." Retrieved August 2018, from <http://crimeandsurveillance.blogspot.com/2005/04/response-to-new-surveillance-cameras.html>
- Lenhart, L. (2013a, July 21). Summer 2013 Information Ethics Blog: Privacy. Retrieved August 2018, from <http://2013infoethics.blogspot.com/2013/07/privacy.html>
- Lenhart, L. (2013b, August 4). Summer 2013 Information Ethics Blog: Course wrap up: Value conflicts and the role of the professional. Retrieved August 2018, from <http://2013infoethics.blogspot.com/2013/08/course-wrap-up-value-conflicts-and-role.html>
- Leo McKinstry. (2009, August 19). Retrieved August 2018, from <http://ironwand.blogspot.com/2009/08/leo-mckinstry.html>
- Marino, J. F. (2014, August 13). FBI Gestapo Blocking Access To This Site Via Google [...]. Retrieved August 2018, from <http://9-11themotherofallblackoperations.blogspot.com/2014/08/the-state-of-new-hampshire-passes-law.html>

- McIntyre, D. (2012, March 15). 2020 Vision secures Trusts trust. Retrieved April 2019, from <http://2020cctv.blogspot.com/2012/03/2020-vision-secures-trusts-trust.html>
- MGT, M. (2005, April 25). EXTERNAL ARTICLE: The screening of America: Crime, cops, and cameras—Surveillance cameras and right to privacy. Retrieved April 2019, from <https://crimeandsurveillance.blogspot.com/>
- News Agency News247WorldPress. (2016, August 26). Country Reports on Terrorism 2015. Retrieved August 2018, from <https://news247worldpressuk.com/2016/08/26/country-reports-on-terrorism-2015/>
- Pad, S. (2010, March 11). Opposing Arguments to the Use of CCTV Security Camera Surveillance. Retrieved August 2018, from <http://freecctvsecuritycamera.blogspot.com/2010/03/opposing-arguments-to-use-of-cctv.html>
- Patrick. (2013, December 14). Patrick's J150 ePortfolio: With modern surveillance techniques and technologies, it is possible to physically go missing but impossible to disappear digitally. Retrieved April 2019, from http://150patrick44.blogspot.com/2013/12/with-modern-surveillance-techniques-and_14.html
- Pax, S. (n.d.). Retrieved April 2018, from <https://salampax.wordpress.com/>
- PeterBDunn. (2014, January 26). ECHELON AND HIGH TECH INTELLIGENCE records for you. Retrieved July 2019, from <https://peterbdunn.wordpress.com/2014/01/26/echelon-and-high-tech-intelligence-records-for-you/>
- privacynewshighlights. (2014, December 1). 16-30 November 2014. Retrieved August 2018, from <https://privacynewshighlights.wordpress.com/2014/12/01/16-30-november-2014/>
- Roger. (n.d.). A force more powerful-freedom. Retrieved July 2019, from <http://libertiesanddenialsof.blogspot.com/>
- Searcher. (2016, February 4). 4. New Postings. Retrieved August 2018, from <http://hospitalsandprivacy.blogspot.com/>
- Severino. (2012, June 29). Calais: Olympic Evictions of Migrant Squats. Retrieved August 2018, from <http://lettheolympicgamesdieintheland.blogspot.com/2012/06/calais-olympic-evictions-of-migrant.html>
- Stephens, N. (2015, March 31). NEV'S BIBLE PROPHECY BLOG FOR THE ANTICHRIST, THE FALSE PROPHET, THE RAPTURE, AND THE END TIMES: BIG BROTHER IS NOT JUST WATCHING YOU - HE IS LISTENING, COLLECTING DATA AND MONITORING EVERYTHING ABOUT YOU!! Retrieved August 2018, from <https://propheticwitnessmovementnewportbranch.blogspot.com/2015/03/big-brother-is-not-just-watching-you-he.html>
- tan, jason. (2004, December 9). A&S Mag, Taipei: IBM Launches Fingerprint Notebook. Retrieved August 2018, from <http://jasontlkas.blogspot.com/2004/12/ibm-launches-fingerprint-notebook.html>
- thinkingshift. (2010, January 17). 2020: Not looking so good. Retrieved August 2019, from <https://thinkingshift.wordpress.com/2010/01/17/the-world-in-2020/>
- Undercover1. (2015, July 2). A Journalists & whistleblowers brief guide to sabotaging gagging laws. Retrieved August 2018, from <https://undercoverinfo.wordpress.com/2015/07/02/a-journalists-whistleblowers-brief-guide-to-sabotaging-gagging-laws/>
- Vanacore, J. (2010, February 27). Cold Class Communications: Internet Privacy- Final Paper. Retrieved June 2018, from <http://coldclasscommunications.blogspot.com/2010/02/internet-privacy-final-paper.html>
- Warnings for America's Future. (n.d.). Warning's for America's Future!: 1,063 Documented Examples of Barack Obama's Lying, Lawbreaking, Corruption, Cronyism, Hypocrisy,

- Waste, Etc. Retrieved July 2019, from <http://warningsforamericasfuture.blogspot.com/2015/05/normal-0-false-false-false-en-us-x-none.html>
- Why Use Burglar Security Alarm 24/7 Monitoring. (2013, July 4). Retrieved April 2019, from <http://eyespyelectronics.blogspot.com/2013/07/why-use-burglar-security-alarm-247.html>
- Who Really Owns and Controls the Military-Industrial Complex and What Are They Doing? – An Extensive Research Report | Stillness in the Storm. (2016, May 11). Retrieved August 2019, from <http://sitsshow.blogspot.com/2016/05/Who-Really-Owns-and-Controls-the-Military-Industrial-Complex-and-What-Are-They-Doing-An-Extensive-Research-Report.html>
- Zeese, K. (2013, December 31). Counter-Insurgency: Stratfor Monitors and Studies Social Movements. Retrieved August 2018, from <https://warriorpublications.wordpress.com/2013/12/31/counter-insurgency-stratfor-monitors-and-studies-social-movements/>

References of cited examples from the Times Digital Archive

- Applying for Jobs. (1997, October 14). *The Times*, p. 42. Retrieved from The Times Digital Archive.
- BBC 1. (1997, June 28). *The Times*, p. 44[S2]+. Retrieved from The Times Digital Archive.
- Bird, S. (2008, July 12). 21/7 leader's fiancée is jailed for disguising him in burka. *The Times*, p. 34. Retrieved from The Times Digital Archive.
- Blunkett, D. (2001, October 4). At times like these, the majority must be protected from the minority. *The Times*, p. 18. Retrieved from The Times Digital Archive.
- Bodkin, W. (1997, September 3). Why paparazzi are beyond the law. *The Times*, p. 23. Retrieved from The Times Digital Archive.
- Bruxelles, S. de. (2008, April 11). Council spies on family trying to ensure a school place. *The Times*, p. 31. Retrieved from The Times Digital Archive.
- Chinese TV turns back the pages. (1986, October 24). *The Times*, p. 10. Retrieved from The Times Digital Archive.
- Connew, P. (1997, September 6). Intrusion and the law. *The Times*, p. 21. Retrieved from The Times Digital Archive.
- Davenport, P. (1986, April 15). Police may seek help of psychologists as stress cases mount. *The Times*, p. 5. Retrieved from The Times Digital Archive.
- Ed Potton. (2008, May 9). Today's TV. In *The Times* (p. 20[S]). Retrieved from The Times Digital Archive.
- Elliott, V., & Duncan, G. (2007, November 14). Deadly flu infects free-range birds. *The Times*, p. [1]+. Retrieved from The Times Digital Archive.
- First Inertial Systems Limited. (1986, July 3). *The Times*, p. 30. Retrieved from The Times Digital Archive.
- Ford, R. (2002, July 4). Compulsory card the key to benefits, NHS and schools. *The Times*, p. 13. Retrieved from The Times Digital Archive.
- Ford, R. (2006, November 2). By 2016, they'll be able to watch you everywhere. *The Times*, p. 6. Retrieved from The Times Digital Archive.
- Ford, R. (2008a, May 31). Do we really need to use these powers to tackle dog fouling? *The Times*, p. 7. Retrieved from The Times Digital Archive.
- Ford, R. (2008b, June 9). ID cards could be used for extra surveillance, MPs fear. *The Times*, p. 11. Retrieved from The Times Digital Archive.
- Ford, R., & O'Neill, S. (2008, February 5). Illegal bug uncovered in second UK prison. *The Times*, p. [1]+. Retrieved from The Times Digital Archive.
- Gibb, F. (1993, January 13). Controls by the state would be "wrong and very dangerous." *The Times*, p. 3. Retrieved from The Times Digital Archive.
- Gibb, F. (1997, January 21). Concessions may save new law from string of challenges. *The Times*, p. 11. Retrieved from The Times Digital Archive.
- Hadley Wood. (1997, May 10). *The Times*, p. 9[S]. Retrieved from The Times Digital Archive.
- Henderson, M. (1997, October 29). Videos of child cruelty branded deceptive. *The Times*, p. 4. Retrieved from The Times Digital Archive.
- Howard, M. (1997, March 12). Home Secretary replies on bugging. *The Times*, p. 19. Retrieved from The Times Digital Archive.
- Index. (2008, April 8). *The Times*, p. [1][S1]. Retrieved from The Times Digital Archive.

- Linklater, M. (1997, August 28). The Princess and our privacy. *The Times*, p. 16. Retrieved from The Times Digital Archive.
- Lorraine Electronics. (1997, February 18). *The Times*, p. 30. Retrieved from The Times Digital Archive.
- Maddox, B. (1997, April 16). How do you protect the unprotectable? *The Times*, p. 26. Retrieved from The Times Digital Archive.
- Miles, A. (2005, June 29). Mission: Identify ID goals. *The Times*, p. 18. Retrieved from The Times Digital Archive.
- Mostrous, A. (2008a, May 31). "I can't believe they were watching my children." *The Times*, p. 6. Retrieved from The Times Digital Archive.
- Mostrous, A. (2008b, May 31). Terror law turns thousands of council officials into ... *The Times*, p. 6+. Retrieved from The Times Digital Archive.
- Mostrous, A. (2008c, July 19). Businessman tells of naked ordeal in knife raid at home. *The Times*, p. 15. Retrieved from The Times Digital Archive.
- Multiple Display Advertising Items. (1986, April 11). *The Times*, p. 22. Retrieved from The Times Digital Archive.
- O'Neill, S., & Coates, S. (2007, May 2). Police "kept in dark" on suspicions about 7/7 ringleader. *The Times*, p. 26. Retrieved from The Times Digital Archive.
- Purves, L. (2002, July 2). Never mind our liberty, consider the practicality. *The Times*, p. 18. Retrieved from The Times Digital Archive.
- Reid, M. (2007, May 31). Nearly all CCTV cameras "are illegal." *The Times*, p. 33. Retrieved from The Times Digital Archive.
- Rentals. (1986, April 28). *The Times*, p. 27. Retrieved from The Times Digital Archive.
- Rifkind, H. (2008, November 21). Big Brother only wants to help. *The Times*, p. 2[S]+. Retrieved from The Times Digital Archive.
- Ruby Electronics Ltd. (1986, March 14). In *The Times* (p. 22). Retrieved from The Times Digital Archive.
- The Lives of Others. (2008, May 31). *The Times*, p. 18. Retrieved from The Times Digital Archive.
- The Times. (2008, April 18). *The Times*, p. 11[S]. Retrieved from The Times Digital Archive.
- The Top Stories. (2006, April 4). *The Times*, p. 2[S1]. Retrieved from The Times Digital Archive.
- The untouchables need electronic eye. (1997, December 6). *The Times*, p. 34. Retrieved from The Times Digital Archive.
- Video captures attack on baby. (1997, October 28). *The Times*, p. 8. Retrieved from The Times Digital Archive.
- Webster, P. (2006, March 17). Fight with Lords over terrorism measures continues. *The Times*, p. 34. Retrieved from The Times Digital Archive.
- Weldon, N. (2007, April 20). The gun is still the guarantor of American freedom [Letter to the editor]. *The Times*, p. 22. Retrieved from The Times Digital Archive.

Appendix A: List of articles included in the S&S Corpus

Article ID	Author(s)	Title
2002-01-1-01	Lyon	Surveillance Studies: Understanding visibility, mobility and the phenetic fix.
2002-01-1-02	Marx	What's new about the "new surveillance"? Classifying for change and continuity
2002-01-1-03	Groombridge	Crime control or crime culture TV?
2002-01-1-04	Fopp	Increasing the potential for gaze, surveillance and normalisation: The transformation of an Australian policy for people who are homeless
2002-01-1-05	Taylor	State surveillance and the right to privacy
2002-01-1-06	Ditton	Hair testing: Just how accurate is it?
2002-01-1-08	Stalder	Privacy is not the antidote to surveillance
2003-01-2-01	Ball	The labours of surveillance.
2003-01-2-02	Thompson	Fantasy Island: a Labour Process critique of the 'age of surveillance'.
2003-01-2-03	Stanton et al.	Information technology, privacy, and power within organizations: A view from Boundary Theory and Social Exchange perspectives
2003-01-2-04	Goold	Public area surveillance and police work: The impact of CCTV on police behaviour and autonomy
2003-01-2-05	Holland	Drug testing in the Australian mining industry
2003-01-2-06	Introna	Workplace surveillance 'is' unethical and unfair
2003-01-2-07	Charlesworth	Privacy, personal information and employment
2003-01-2-08	Beck	Working for Them.
2003-01-2-09	Searle	Organizational justice in e-recruiting: Issues and controversies
2003-01-3-01	Wood	Foucault and panopticism revisited
2003-01-3-02	Elden	Plague, panopticon, police
2003-01-3-03	Yar	Panoptic power and the pathologisation of vision: Critical reflections on the Foucauldian Thesis.
2003-01-3-04	Vaz et al.	Types of Self-Surveillance: From abnormality to individuals 'at risk'.
2003-01-3-05	Koskela	'Cam Era' — the contemporary urban Panopticon.
2003-01-3-06	Gray	Urban Surveillance and Panopticism: will we recognize the facial recognition society?
2003-01-3-07	Mann et al.	Sousveillance: Inventing and using wearable computing devices for data collection in surveillance environments
2003-01-3-09	Mann et al.	PanopDecon: Deconstructing, decontaminating, and decontextualizing panopticism in the postcyborg era
2003-01-3-10	Hier	Probing the surveillant assemblage: On the dialectics of surveillance practices as processes of social control.
2003-01-3-11	Lianos	Social Control After Foucault / Le Contrôle Social après Foucault.
2003-01-4-01	Bennett et al.	Surveillance and mobilities
2003-01-4-02	Arvidsson	On the 'Pre-History of The Panoptic Sort': Mobility in market research
2003-01-4-03	Curry	The profiler's question and the treacherous traveler: Narratives of belonging in commercial aviation
2003-01-4-04	Adey	Secured and Sorted Mobilities: Examples from the Airport.
2003-01-4-05	Sweeny	Between 'Devil' and Detournement: Embodied acts as methods of critical inquiry in educational spaces
2003-01-4-06	Fotel et al.	The surveillance of children's mobility
2003-01-4-07	Gow et al.	Prepaid mobile phone service and the anonymous caller: Considering Wireless E9-1-1 in Canada.
2003-01-4-08	Green et al.	'A spy in your pocket'? The regulation of mobile data in the UK
2004-02-1-01	Williams et al.	Circuits of surveillance
2004-02-1-02	Los	The technologies of total domination

2004-02-1-03	Whitty	Should filtering software be utilised in the workplace? Australian employees' attitudes towards internet usage and surveillance of the internet in the workplace.
2004-02-1-04	Yacobi	In-between surveillance and spatial protest: The production of space of the 'Mixed City' of Lod
2004-02-1-05	Campbell	Technologies of suspicion: Coercion and compassion in post-disciplinary surveillance regimes
2004-02-1-06	Myrick	Escape from the carceral: Writing by American prisoners, 1895-1916
2004-02-2_3-01	Norris et al.	The growth of CCTV: A global perspective on the international diffusion of video surveillance in publicly accessible space.
2004-02-2_3-02	Cameron	CCTV and (In)dividuation
2004-02-2_3-03	Klauser	A comparison of the impact of protective and preservative video surveillance on urban territoriality: The case of Switzerland
2004-02-2_3-04	Muller et al.	Zurich Main Railway Station: A typology of public CCTV Systems
2004-02-2_3-05	Introna et al.	Picturing Algorithmic Surveillance: The Politics of Facial Recognition Systems
2004-02-2_3-06	Koskela	Webcams, TV shows and mobile phones: Empowering exhibitionism
2004-02-2_3-07	Gras	The Legal Regulation of CCTV in Europe
2004-02-2_3-08	Webster	The Diffusion, Regulation and Governance of Closed-Circuit Television in the UK
2004-02-2_3-09	Fussey	New labour and new surveillance: Theoretical and political ramifications of CCTV implementation in the UK
2004-02-2_3-10	Gallagher	CCTV and human rights: The fish and the bicycle? An Examination of Peck V. United Kingdom (2003) 36 E.H.R.R. 41
2004-02-2_3-11	Coleman	Reclaiming the streets: Closed circuit television, neoliberalism and the mystification of social divisions in liverpool, uk
2004-02-2_3-12	Sutton et al.	Open-Street CCTV in Australia: The politics of resistance and expansion
2004-02-2_3-13	Helten et al.	Reactive attention: Video Surveillance in Berlin shopping malls
2004-02-2_3-14	Lomell	Targeting the Unwanted: Video surveillance and categorical exclusion in Oslo, Norway
2004-02-2_3-15	Martinals et al.	Social Aspects of CCTV in France: The case of the city centre of Lyons
2004-02-2_3-16	Smith	Behind the screens: Examining constructions of deviance and informal practices among CCTV control room operators in the UK
2004-02-2_3-17	Saetnan et al.	Controlling CCTV in Public Spaces: Is Privacy the (Only) Issue? Reflections on Norwegian and Danish Observations.
2004-02-2_3-18	Ruegg et al.	CCTV, risk management and regulation mechanisms in publicly-used places: A discussion based on Swiss examples
2004-02-2_3-19	Cole	Signage and Surveillance: Interrogating the Textual Context of CCTV in the UK
2004-02-2_3-20	Jorgensen	The apple of the eye: Parents' use of webcams in a Danish day nursery
2004-02-2_3-21	Kammerer	Video surveillance in Hollywood movies
2004-02-4-01	Wood	People watching people
2004-02-4-02	Andrejevic	The work of watching one another: Lateral surveillance, risk, and governance
2004-02-4-03	Zurawski	"I know where you live!" – Aspects of watching, surveillance and social control in a conflict zone (Northern Ireland)
2004-02-4-04	Lister	Controlling letting arrangements? Landlords and surveillance in the private rented sector
2004-02-4-05	Wakefield	The public surveillance functions of private security
2004-02-4-06	Dubbeld	Protecting personal data in camera surveillance practices
2004-02-4-07	Patton	An exploration of the external validity of self-report amongst arrestees
2004-02-4-08	Jones	A tagging tale: The work of the monitoring officer, electronically monitoring offenders in England and Wales
2004-02-4-09	Jones	Visible rights: Watching out for women
2004-02-4-10	Mann	People watching people watchers: 'The Law Enforcement Company' for watching over those who come to see and be seen on the 'Urban Beach'
2005-03-1-01	Simon	The return of panopticism: Supervision, subjection and the new surveillance

2005-03-1-02	Finn	Photographing fingerprints: Data collection and state surveillance
2005-03-1-03	Gibbs et al.	Data quality, database fragmentation and information privacy
2005-03-1-04	Howard et al.	Digital technology and the market for political surveillance
2005-03-1-05	Lee	Internet control and auto-regulation in Singapore
2005-03-1-06	Powell et al.	Surveillance and Morality: Revisiting the Education Reform Act (1988) in the United Kingdom
2005-03-1-07	Parker et al.	Mutual Obligation? Regulating by supervision and surveillance in Australian income support policy
2005-03-2_3-01	Ball et al.	Doing Surveillance Studies.
2005-03-2_3-02	Kemple et al.	Observing the observers: Researching surveillance and counter-surveillance on 'Skid Row'
2005-03-2_3-03	Walby	Institutional ethnography and surveillance studies: An outline for inquiry
2005-03-2_3-04	Haggerty et al.	The public politics of opinion research on surveillance and privacy
2005-03-2_3-05	Samatas	Studying surveillance in Greece: Methodological and other problems related to an authoritarian surveillance culture.
2005-03-2_3-06	Wright	The ECHELON Trail: An illegal vision
2005-03-2_3-07	Albrechtslund et al.	The plays and arts of surveillance: Studying surveillance as entertainment
2005-03-2_3-08	Marks	Imagining surveillance: Utopian visions and surveillance studies
2005-03-2_3-09	Sweeny	Para-sights: Multiplied Perspectives on surveillance research in art educational spaces
2005-03-2_3-10	Beaumont	Using CCTV to study visitors in The New Art Gallery, Walsall, UK.
2006-04-1_2-01	Genosko et al.	Administrative surveillance of alcohol consumption in Ontario, Canada: Pre electronic technologies of control.
2006-04-1_2-02	Walby	Little England? The rise of open-street Closed-Circuit Television surveillance in Canada.
2006-04-1_2-03	Flint	Surveillance and exclusion practices in the governance of access to shopping centres on periphery estates in the UK.
2006-04-1_2-04	Dawson	The impact of institutional surveillance technologies on student behaviour.
2006-04-1_2-05	Hempel	In the eye of the beholder? Representations of video surveillance in German public television
2006-04-1_2-06	Harris	The Omniscient Eye: Satellite Imagery, "Battlespace Awareness," and the Structures of the Imperial Gaze.
2006-04-1_2-07	Perault	Community life in Colombia under the surveillance of extreme right paramilitary organizations.
2006-04-1_2-08	Wright	Sub-lethal vision: varieties of military surveillance technology
2007-04-3-01	Monahan et al.	Somatic Surveillance: Corporeal control through information networks
2007-04-3-02	Minnaar	The implementation and impact of crime prevention / crime control open street Closed-Circuit Television surveillance in South African Central Business Districts
2007-04-3-03	Bloss	Escalating U.S. police surveillance after 9/11: An examination of causes and effects
2007-04-3-04	Fussey	An interrupted transmission? Processes of CCTV implementation and the impact of human agency.
2007-04-3-05	Marks	Drug detection dogs and the growth of olfactory surveillance: Beyond the rule of law?
2007-04-3-06	Stenson	Harcourt's Against Prediction
2007-04-3-07	Webster	Tator and Henry's Racial Profiling in Canada
2007-04-3-08	Gates	Turow's Niche Envy
2007-04-4-01	Smith	Exploring Relations between watchers and watched in control(led) systems: Strategies and tactics
2007-04-4-02	Paterson	Street-level surveillance: Human agency and the electronic monitoring of offenders
2007-04-4-03	Boa	Privacy outside the castle: Surveillance technologies and reasonable expectations of privacy in Canadian judicial reasoning
2007-04-4-04	Albrechtslund	Balkin and Noveck's The State of Play

2007-04-4-05	Maxwell	Kackman's Citizen Spy
2007-04-4-06	Andrejevic	Virilio's The Original Accident
2007-04-4-07	Genosko	Fischer's Digital Shock
2008-05-1-01	Harper	The politics of paranoia: Paranoid positioning and conspiratorial narratives in the surveillance society
2008-05-1-02	Carroll-Mayer et al.	CCTV identity management and implications for criminal justice: Some considerations
2008-05-1-03	Zurawski et al.	Crime, maps and meaning: Views from a survey on safety and CCTV in Germany.
2008-05-1-04	Groombridge	Stars of CCTV? How the Home Office wasted millions - A radical 'Treasury/Audit Commission' view.
2008-05-1-05	Sewell	Du Gay's The Values of Bureaucracy
2008-05-1-06	Dahl	Neyland's Privacy, Surveillance and Public Trust
2008-05-1-07	Saetnan	Schweber's Disciplining Statistics
2008-05-1-08	Ball	Shulman's From Hire to Liar
2008-05-1-09	Sheptycki	Wall's Cybercrime
2008-05-1-10	Johnson	New edition of Diffie and Landau's Privacy on the Line
2008-05-2-01	Amoore et al.	Smart borders and mobilities: Spaces, zones, enclosures
2008-05-2-02	Ceyhan	Technologization of security: Management of uncertainty and risk in the age of biometrics
2008-05-2-03	Wilson et al.	Surveillance, risk and preemption on the Australian border
2008-05-2-04	Cote-Boucher	The diffuse border: Intelligence-sharing, control and confinement along Canada's smart border
2008-05-2-05	de Lint	The security double take: The political, simulation and the border
2008-05-2-06	Codourey	Mobile identities and the socio-spatial relations of air travel
2008-05-2-07	Zimmer	Lyon's Theorizing Surveillance
2008-05-2-08	Fonio	Monahan's Surveillance and Security
2008-05-2-09	Fisher	Poster's Information Please
2008-05-2-10	Brighenti	González-Crussi's On Seeing
2008-05-2-11	Mair	Shih's Visuality and Identity
2008-05-2-12	Wills	Reinhardt, Edwards and Duganne's Beautiful Suffering, and Apel and Smith's Lynching Photographs
2008-05-3-01	Monahan	Surveillance and inequality
2008-05-3-02	Willse	"Universal Data Elements," or the biopolitical life of homeless populations
2008-05-3-03	Kenner	Securing the elderly body: Dementia, surveillance, and the politics of "aging in place"
2008-05-3-04	Kanashiro	Surveillance cameras in Brazil: Exclusion, mobility regulation, and the new meanings of security
2008-05-3-05	Doherty et al.	Homelessness and exclusion: Regulating public space in European Cities
2008-05-3-06	Pallitro et al.	Theorizing cross-border mobility: Surveillance, security and identity
2008-05-3-07	Wiebe	Re-thinking citizenship: (Un)Healthy bodies and the Canadian Border
2008-05-3-08	Hagmann	Ceyhan's Identifier et Surveiller
2008-05-3-09	Stuart	Posner and Vermeule's Terror in the Balance
2008-05-3-10	Szekely	Henriques' Corporate Truth
2008-05-3-11	Wall	Balkin et al.'s Cybercrime
2009-06-1-01	Murakami Wood	A new 'baroque arsenal'? Surveillance in a global recession
2009-06-1-02	Williams et al.	Police filming English streets in 1935: The limits of mediated identification
2009-06-1-03	Webster	CCTV policy in the UK: Reconsidering the evidence base
2009-06-1-04	Klauser	'Lost' Surveillance Studies: A critical review of French work on CCTV
2009-06-1-05	Douglas	Disappearing Citizenship: Surveillance and the state of exception
2009-06-1-06	Kammerer	Police use of public video surveillance in Germany 1956: management of traffic, repression of flows, persuasion of offenders
2009-06-1-07	Nellis	"I am not a number!" David Davis, The Prisoner and the critique of surveillance

2009-06-1-08	Murakami Wood	Situating surveillance studies. Sean Hier and Josh Greenberg's The Surveillance Studies Reader, and David Lyon's Surveillance Studies: an Overview.
2009-06-1-09	Arteaga Botello	Mattelart's La globalisation de la surveillance
2009-06-1-10	de Lint	Salter's Politics at the Airport
2009-06-1-11	Frois	Samatas' Surveillance in Greece
2009-06-1-13	Hayman	Rajaram and Grundy-Warr's Borderscapes
2009-06-1-14	Koskela	Rule's Privacy in Peril
2009-06-1-15	Leman-Langlois	Deflem's Surveillance and Governance
2009-06-1-16	Lippert	Andrejevic's iSpy
2009-06-1-17	Sealy	Biber's Captive Images
2009-06-1-18	Silverstone	Rigakos' Nightclub
2009-06-1-19	Steeves	Friedman's Guarding Life's Dark Secrets
2009-06-1-20	Thompson	Webb's Illusions of Security
2009-06-1-21	Todd	Hawk et al.'s Small Tech
2009-06-1-22	Wark	Schmeidel's STASI
2009-06-1-23	Yar	Deibert et al.'s Access Denied
2009-06-2-01	Earle et al.	Health, medicine and surveillance in the 21st century
2009-06-2-02	French	Woven of war-time fabrics: The globalization of public health surveillance
2009-06-2-03	Bauer et al.	Observing the others, watching over oneself: Themes of medical surveillance in post-panoptic society
2009-06-2-04	Wiebe	Producing bodies and borders: A review of immigrant medical examinations in Canada
2009-06-2-05	Day	Does my bum look big in this? Reconsidering anorexia nervosa within the culture context of 20th century Australia
2009-06-2-06	Bell	@ the doctor's office': Pro-anorexia and the medical gaze
2009-06-2-07	Rich et al.	Prosthetic Surveillance: The medical governance of healthy bodies in cyberspace
2009-06-2-08	French	Wald's Contagious
2009-06-2-09	Gerlach	Roof's The Poetics of DNA
2009-06-2-10	Mitchell	Miah and Rich's The Medicalization of Cyberspace
2009-06-2-11	Sluggett	Nadesan's Governmentality, Biopower and Everyday Life
2009-06-2-12	Albrechtslund	Rettberg's Blogging
2009-06-2-13	Bogard	Urry's Mobilities
2009-06-2-14	Boyle	Graham's Cities, War and Terrorism
2009-06-2-15	Kenner	Solove's The Digital Person
2009-06-2-16	Nellis	Mullan's Anonymity
2009-06-3-01	Fernandez et al.	Is resistance futile? Thoughts on resisting surveillance
2009-06-3-02	Bell	Surveillance is sexy
2009-06-3-03	Martin et al.	Understanding resistance to digital surveillance: Towards a multi-disciplinary, multi-actor framework
2009-06-3-04	Introna et al.	Networks and resistance: Investigating online advocacy networks as a modality for resisting state surveillance
2009-06-3-05	Wells et al.	Individualism and identity: Resistance to speed cameras in the UK
2009-06-3-06	Sanchez	Facebook Feeding Frenzy: Resistance-through-distance and resistance-through-persistence in the societied network
2009-06-3-07	Marx	A tack in the shoe and taking off the shoe neutralization and counter-neutralization dynamics
2009-06-3-09	Gandy	What the US can learn from the UK about the protection of privacy
2009-06-3-10	Hayles	Waking up to the surveillance society
2009-06-3-11	Aas	Surveillance: Citizens and the state
2009-06-3-12	Andrejevic	Control over personal information in the database era
2009-06-3-13	Bajc	Solove, Daniel J. 2008. Understanding Privacy. Cambridge and London: Harvard University Press.
2009-06-3-14	Boa	Roberts, Alasdair. 2006. Blacked Out: Government Secrecy in the Information Age. Cambridge, MA: Cambridge University Press.

2009-06-3-15	Dodge	Wright, David, Serge Gutwirth, Michael Friedewald, Elena Vildjiounaite and Yves Punie (eds.) 2008. <i>Safeguards in a World of Ambient Intelligence</i> . London: Springer.
2009-06-3-16	Molnar	Franko Aas, Katja, Helene Oppen Gudnhus and Heidi Mork Lomell 2008. <i>Technologies of Insecurity: The Surveillance of Everyday Life</i> . New York: Routledge-Cavendish.
2009-06-3-17	Morrison	Laidler, Keith 2008. <i>Surveillance Unlimited: How We've Become the Most Watched People on Earth</i> . Cambridge: Icon Books.
2009-06-3-18	Murphy	Reiner, Robert 2007. <i>Law and Order: An Honest Citizen's Guide to Crime Control</i> . London: Polity Press.
2009-06-3-19	Nellis	Doctorow, Cory. 2008. <i>Little Brother</i> . London: Harper Voyager or New York: Tor/Tom Doherty Associates.
2009-06-3-20	Newman	Bennett, Colin. 2008. <i>The Privacy Advocates: Resisting the Spread of Surveillance</i> . Cambridge: MIT Press.
2009-06-3-21	Neyland	Solove, Daniel. 2007. <i>The Future of Reputation: Gossip, Rumour, and Privacy on the Internet</i> . London: Yale University Press.
2009-06-3-22	Regan	Sofsky, Wolfgang. 2008. <i>Privacy: A Manifesto</i> . Princeton, NJ: Princeton University Press. Translated by Steven Rendall.
2009-06-3-23	Spink D'Souza	Freeman, D., And J. Freeman 2006. <i>Paranoia: The Twenty-First Century Fear</i> . Oxford University Press.
2009-06-4-01	Ball et al.	Surveillance studies needs gender and sexuality
2009-06-4-02	Beauchamp	Artful concealment and strategic visibility: Transgender bodies and U.S. State surveillance after 9/11
2009-06-4-03	Walby	Ottawa's National Capital Commission Conservation Officers and the Policing of Public Park Sex
2009-06-4-04	Conrad	Surveillance, gender, and the virtual body in the information age
2009-06-4-05	Corones et al.	En-gendered surveillance: Women on the edge of a watched cervix
2009-06-4-07	Raab et al.	Putting surveillance on the political agenda – A short defence of surveillance, citizens and the state
2009-06-4-08	Devers	Review of Hassan, Robert. 2008. <i>The Information Society</i> . Cambridge: Polity Press.
2009-06-4-09	Martin	Review of Klein, Mark. 2009. <i>Wiring Up the Big Brother Machine... And Fighting It</i> . Charleston, South Carolina: BookSurge
2009-06-4-10	Mayer-Schönberger	Review of Newman, Abraham L. 2008. <i>Protectors of Privacy: Regulating Personal Data in the Global Economy</i> . Ithaca: Cornell University Press.
2009-06-4-12	Monahan	Review of Fosher, Kerry B. 2009. <i>Under Construction: Making Homeland Security at the Local Level</i> . Chicago: University of Chicago Press
2009-06-4-13	Staples	Groebner, Valentin. 2007. <i>Who Are You? Identification, Deception, and Surveillance in Early Modern Europe</i> . Translated by Mark Kyburz and John Peck. New York: Zone Books
2009-06-4-14	Zurawski	Review of Rule, James B. (eds) 2008. <i>Global Privacy Protection. The First Generation</i> . Cheltenham, UK; Northampton, MA, USA: Edward Elgar
2009-07-1-01	Murakami Wood	Questions of surveillance
2009-07-1-02	Guzik	Discrimination by design: Predictive data mining as security practice in the United States' 'war on terrorism'
2009-07-1-03	Ikebuchi Ketchell	Carceral ambivalence: Japanese Canadian 'internment' and the Sugar Beet Programme during WWII
2009-07-1-04	Holm	Conspiracy theorizing surveillance: Considering modalities of paranoia and conspiracy in surveillance studies
2009-07-1-05	Gad et al.	Situated surveillance: an ethnographic study of fisheries inspection in Denmark
2009-07-1-06	O'Byrne et al.	Public health STI/HIV surveillance: Exploring the society of control
2009-07-1-07	Hier et al.	David Murakami Wood's 'Situating Surveillance Studies'
2009-07-1-08	Murakami Wood	An olive branch to Hier and Greenberg
2009-07-1-09	Firmino et al.	Book Review of "The Electronic City"
2009-07-1-10	Ellerbrok	Book Review of "Edited Clean Version"
2009-07-1-11	Feng	Book Review of "Private and Confidential?"

2009-07-1-12	Pridmore	Book Review of "Headspace"
2009-07-1-13	Sriraman	Book Review of "Playing the Identity Card"
2010-07-2-01	McGrath et al.	Editorial: Surveillance, performance and new media
2010-07-2-09	Brighenti	Artveillance: At the crossroads of art and surveillance
2010-07-3_4-01	Steeves et al.	Surveillance, children and childhood
2010-07-3_4-02	Marx et al.	From the beginning: Children as subjects and agents of surveillance
2010-07-3_4-03	Henderson et al.	A new state of surveillance? Applying Michel Foucault to Modern Motherhood
2010-07-3_4-04	Sparrman et al.	Visual documentation as a normalizing practice: a new discourse of visibility in preschool
2010-07-3_4-05	Gallagher	Are schools panoptic?
2010-07-3_4-06	McCahill et al.	The social impact of surveillance in three UK schools: Angels, devils and teen mums
2010-07-3_4-07	McIntosh et al.	'You don't have to be watched to make your toast': Surveillance and food practices within residential care
2010-07-3_4-08	Wrenall	Surveillance and child protection: De-mystifying the Trojan Horse
2010-07-3_4-09	Osmond	Anti-social behaviour and its surveillant inter-assemblage
2010-07-3_4-10	Rooney	Trusting Children: How do surveillance technologies alter a child's experience of trust, risk and responsibility?
2010-07-3_4-11	Ragnedda	Review of Monahan and Torres' (eds.) Schools Under Surveillance
2010-08-1-01	Braverman	Governing with clean hands: Automated public toilets and sanitary surveillance
2010-08-1-02	Nunn	'Wanna still nine hard?': Exploring mechanisms of police bias in the translation and interpretation of wiretap conversations
2010-08-1-03	Newkirk	The Rise of the Fusion-Intelligence Complex: A critique of political surveillance after 9/11
2010-08-1-04	Waiton	The Politics of Surveillance: Big Brother on Prozac
2010-08-1-05	Desai	Violence and surveillance: Some unintended consequences of CCTV monitoring within mental health hospital wards
2010-08-1-06	Hiranandani	Under-explored threats to privacy: See-Through-Wall Technologies and Electro-Magnetic Radiations
2010-08-1-07	Martin	Review of Klein's Wiring Up the Big Brother Machine...And Fighting It
2010-08-1-08	McDowell	Review of Clark's The Road to Big Brother
2010-08-1-09	Bernhard	Review of Chertoff's Homeland Security: Assessing the First Five Years
2010-08-2-01	Monahan et al.	Surveillance and empowerment
2010-08-2-02	Walsh	From border control to border care: The political and ethical potential of surveillance
2010-08-2-03	Shilton	Participatory sensing: Building empowering surveillance
2010-08-2-04	Regan et al.	Kids R Us: Online Social Networking and the Potential for Empowerment
2010-08-2-05	Wilson et al.	Video activism and the ambiguities of counter-surveillance
2010-08-2-06	Skoric et al.	Online shaming in the Asian context: Community empowerment or civic vigilantism?
2010-08-2-07	Ellerbrok	Empowerment: analyzing technologies of multiple variable visibility
2010-08-2-08	Ottinger	Constructing empowerment through interpretations of environmental surveillance data
2010-08-2-09	Albrechtslund et al.	Empowering residents: A theoretical framework for negotiating surveillance technologies
2010-08-2-10	Ball	Review of Thompson and Genosko's Punched Drunk
2010-08-2-11	Barnard-Wills	Review of Lyon's Identifying Citizens
2010-08-2-12	Datta	Review of Kelly's The Political Philosophy of Michel Foucault
2010-08-2-13	Ferenbok	Review of Pegler Gordon's In Sight of America
2010-08-2-14	Kerr	Review of Mayer-Shonberger's Delete
2010-08-2-15	Klauser	Review of Andrzejewski's Building Power
2010-08-2-16	Murakami Wood	Review of McCoy's Policing America's Empire
2011-08-3-01	Pridmore et al.	Marketing and the Rise of Commercial Consumer Surveillance
2011-08-3-02	Andrejevic	Surveillance and Alienation in the Online Economy

2011-08-3-03	Fuchs	Web 2.0, presumption, and surveillance
2011-08-3-04	McStay	Profiling Phorm: an autopoietic approach to the audience-as-commodity
2011-08-3-05	Manzerolle et al.	Consumer Databases, Neoliberalism, and the Commercial Mediation of Identity: A Medium Theory Analysis
2011-08-3-06	Palmås	Predicting what you'll do tomorrow: Panspectric surveillance and the contemporary corporation
2011-08-3-07	Murià et al.	Shopping and Working in the Borderlands: Enforcement, Surveillance and Marketing in Tijuana, Mexico
2011-08-3-08	Jurgenson	Review of Timoner's We Live in Public
2011-08-3-09	Barnard-Wills	Review of Gandy's Coming to Terms with Chance
2011-08-3-10	Pentecost	Review of Engle's Seeing Ghosts
2011-08-3-11	Walby	Review of Jensen's Cautious Beginnings
2011-08-3-12	Fonio	Review of Marton's Enemies of the People
2011-08-3-13	McLane	Review of Mountz's Seeking Asylum
2011-08-3-14	van Brakel	Review of Nelson and Garey's Who's Watching?
2011-08-3-15	Wilson	Review of Natapoff's Snitching
2011-08-4-01	Sperling	The politics of transparency and surveillance in post-reunification Germany
2011-08-4-02	Finn	Surveillant staring: Race and the everyday surveillance of South Asian women after 9/11
2011-08-4-03	Zimmer	Surveillance cinema: Narrative between technology and politics
2011-08-4-04	Taekke	Digital panopticism and organizational power
2011-08-4-05	Taylor	A conceptual legal framework for privacy, accountability and transparency in visual surveillance systems
2011-08-4-06	Svedberg Helgesson	Public-private partners against crime: Governance, surveillance and the limits of corporate accountability
2011-08-4-07	Bennett	In defence of privacy: The concept and the regime
2011-08-4-08	Regan	Response to Bennett: Also in defence of privacy
2011-08-4-09	Gilliom	A response to Bennett's 'In defence of privacy'
2011-08-4-10	boyd	Dear Voyeur, Meet Flâneur... Sincerely, Social Media
2011-08-4-11	Stalder	Autonomy beyond Privacy? A Rejoinder to Colin Bennett
2011-08-4-12	Bennett	In further defence of privacy...
2011-08-4-13	Cousineau	The surveillant simulation of war: entertainment and surveillance in the 21st century
2011-08-4-14	Galdon Clavell	The political economy of surveillance in the (wannabe) global city
2011-08-4-15	Stedmon	The camera never lies, or does it? The dangers of taking CCTV surveillance at face value
2011-08-4-16	Muller	Review of Pugliese's Biometrics: Bodies, Technologies, Biopolitics
2011-08-4-17	Phillips	Review of Monahan's Surveillance in Time of Insecurity
2011-08-4-18	Leman-Langlois	Review of Hier's Panoptic Dreams
2011-08-4-19	Bennett	Review of Nissenbaum's Privacy in Context
2011-08-4-20	Fussey	Review of Haggerty and Samatas's Surveillance and Democracy
2011-08-4-21	Trottier	Review of Weir & Mykhalovskiy's Global Public Health Vigilance
2011-08-4-22	Spiller	Review of Deane's Big Brother Watch
2011-09-1_2-01	Monahan et al.	Surveillance impediments: Recognizing obduracy with the deployment of hospital information systems
2011-09-1_2-02	Trottier	Mutual transparency or mundane transgressions? Institutional Creeping on Facebook
2011-09-1_2-03	Wang et al.	Deviance and control in communities with perfect surveillance – The case of Second Life
2011-09-1_2-04	Maki	Neoliberal deviants and surveillance: Welfare recipients under the watchful eye of Ontario Works
2011-09-1_2-05	Çavlin Bozbeyoğlu	Citizenship rights in a surveillance society: The case of the electronic ID card in Turkey
2011-09-1_2-06	Fonio	Surveillance under Mussolini's Regime
2011-09-1_2-07	Bozzini	Low-tech surveillance and the despotic state in Eritrea

2011-09-1_2-08	Porter et al.	Mobility, surveillance and control of children and young people in the everyday: Perspectives from sub-Saharan Africa
2011-09-1_2-09	Ema et al.	How far can child surveillance go?: Assessing the parental perceptions of an RFID child monitoring system in Japan
2011-09-1_2-10	Smith et al.	Playing with surveillance: The design of a mock RFID-based identification infrastructure for public engagement
2011-09-1_2-11	Thiessen	Being watched watching watchers watch: Determining the Digitized future while profitably modulating preemption (at the airport)
2011-09-1_2-12	Gschrey	Borderlines. Surveillance, Identification and Artistic Explorations along European Borders
2011-09-1_2-13	Liu	The Wall, the Window and the Alcove: Visualizing Privacy
2011-09-1_2-14	Steiner et al.	Living Behind Glass Facades: Surveillance Culture and New Architecture
2011-09-1_2-15	Bright	Building Biometrics: Knowledge Construction in the Democratic Control of Surveillance Technology
2011-09-1_2-16	Marx	'Cough a little upon entering': Some reflections on the history of surreptitious spectatorship - An extended review of: Locke's Eavesdropping: An Intimate History.
2012-09-3-01	Lippert et al.	New Urban Surveillance: Technology, Mobility, and Diversity in 21st Century Cities
2012-09-3-02	Muir	'Control Space?': Cinematic representations of surveillance space between discipline and control'
2012-09-3-03	van Liempt et al.	Urban surveillance and the struggle between safe and exciting nightlife districts
2012-09-3-04	Palmer et al.	ID scanning, The media and the politics of urban surveillance in an Australian Regional City
2012-09-3-05	Lippert et al.	Municipal corporate security and the intensification of urban surveillance
2012-09-3-06	Graham	Digital medieval
2012-09-3-07	Monaghan	Review of Graham's Cities Under Siege: The New Military Urbanism
2012-09-3-08	Zurawski	Review of Bennett and Haggerty's Security Games: Surveillance and Control at Mega-Events
2012-09-3-09	Jeandesboz	Review of Basaran's Security, Law and Borders
2012-09-3-10	Staples	Review of Robertson's The Passport in America: The History of a Document
2012-09-3-11	Gates	Review of Nelson's America Identified
2012-09-4-01	Bennett et al.	Introduction to cyber-surveillance
2012-09-4-02	Mueller et al.	Policing the network: Using DPI for copyright enforcement
2012-09-4-03	Meyer et al.	Surveillance and regulating code: An Analysis of graduated response in France
2012-09-4-04	Marwick	The public domain: Surveillance in everyday life
2012-09-4-05	Draper	Group Power: Discourses of Consumer Power and Surveillance in Group Buying Websites
2012-09-4-06	Martin	Second Life SURVEILLANCE: Power to the people or virtual surveillance society?
2012-09-4-07	Farrall	Online Collectivism, Individualism and Anonymity in East Asia
2012-09-4-08	Leistert	Resistance against Cyber-Surveillance within Social Movements and how Surveillance Adapts
2012-09-4-09	Gilliom	Review of Nippert-Eng's Islands of Privacy
2012-09-4-10	Milbrandt	Review of Mirzoeff's The Right to Look
2012-09-4-11	Sandhu	Review of Sussman's The Propaganda Society
2012-10-1-01	Firmino et al.	Understanding the Sociotechnical Networks of Surveillance Practices in Latin America
2012-10-1-02	Botello Arteaga	Surveillance Studies: An Agenda for Latin America
2012-10-1-03	Kanashiro et al.	The new Brazilian identification system: Unique features of a general transformation
2012-10-1-04	Firmino et al.	Eyes of glass: Watching the Watchers in the monitoring of public places in Curitiba, Brazil
2012-10-1-05	Vélez	Insecure identities. The approval of a biometric ID Card in Mexico

2012-10-1-06	de Vasconcelos Cardoso	The paradox of caught-in-the-act surveillance scenes: Dilemmas of police video surveillance in Rio de Janeiro
2012-10-1-07	Lissofsky et al.	Images of suspicion: Surveillance photos in the Brazilian Political Police Archives
2012-10-1-09	Mitchell	Review of Arribas-Ayllon, Sarangi and Clarke's Genetic Testing
2012-10-1-10	O'Connor	Review of Coleman and McCahill's Surveillance and Crime
2012-10-1-11	Bisaillon	Review of Gates' Our Biometric Future
2012-10-1-12	Bull	Review of Carter's Analysing Police Interviews
2012-10-1-13	Coleman	Review of Doyle, Lippert and Lyon's Eyes Everywhere
2012-10-1-14	Smith	Review of Bajc and de Lint's Security and Everyday Life
2012-10-2-01	Mason et al.	Surveillance Studies and Violence Against Women
2012-10-2-02	Braverman	Zooveillance: Foucault Goes to the Zoo
2012-10-2-03	Finn	Surveillance Studies and Visual Art: An Examination of Jill Magid's Evidence Locker
2012-10-2-04	Anderson	The Spy in the Cab: The Use and Abuse of Taxicab Cameras in San Francisco
2012-10-2-05	Timan et al.	Mobile cameras as new technologies of surveillance? How citizens experience the use of mobile cameras in public nightscapes
2012-10-2-06	Lohne	The Norwegian Data Inspectorate: Between governance and resistance
2012-10-2-07	Goldsmith	Review of Strangelove's Watching YouTube
2012-10-2-08	Superle	Review of Dubrofsky's The Surveillance of Women on Reality Television
2012-10-2-09	Trottier	Review of Allmer's Towards a Critical Theory of Surveillance in Informational Capitalism
2012-10-3_4-01	Barnard-Wills et al.	Invisible Surveillance in Visual Art
2012-10-3_4-02	Patel	Surveillance, Suspicion and Stigma: Brown Bodies in a Terror-Panic Climate
2012-10-3_4-03	Reeves	If You See Something, Say Something: Lateral Surveillance and the Uses of Responsibility
2012-10-3_4-04	Tewksbury	Crowdsourcing homeland security: The Texas Virtual BorderWatch and participatory citizenship
2012-10-3_4-05	Luther et al.	Perspectives on privacy, information technology, and company/governmental surveillance in Japan
2012-10-3_4-06	Backman	Mandatory criminal record checks in Sweden: Scandals and function creep
2012-10-3_4-07	Due et al.	Surveillance, security and violence in a mental health ward: An ethnographic case- study of an Australian purpose-built unit
2012-10-3_4-08	Manley et al.	Disciplinary power, the oligopticon and rhizomatic surveillance in elite sports academies
2012-10-3_4-09	Rapoport	The home under surveillance: A tripartite assemblage
2012-10-3_4-10	Huey	The Review of Privacy Online: Three Questions
2012-10-3_4-11	Vonn	Talkin' About A Revolution
2012-10-3_4-12	Whitaker	The curious tale of the dog that hasn't barked (yet)
2012-10-3_4-13	Rosenzweig	Whither privacy?
2012-10-3_4-14	Boyd	Networked privacy
2012-10-3_4-15	Margulis et al.	Caveat Emptor: A perspective on consumer privacy online
2012-10-3_4-16	Rauhofer	Future-proofing privacy: Time for an ethical introspection?
2012-10-3_4-17	Wright	Review of Ball, Haggerty and Lyon's Handbook of Surveillance Studies
2012-10-3_4-18	Vite Pérez	Review of Bruno, Kanashiro and Firmino's Vigilância e Visibilidade: Espaço, tecnologia e identificação
2012-10-3_4-19	Haggerty	Review of Braverman's Zooland
2013-11-1_2-01	Ball et al.	Political economies of surveillance
2013-11-1_2-02	Pallitto	Bargaining with the machine: A framework for describing encounters with surveillance technologies
2013-11-1_2-03	Mann et al.	New media and the power politics of sousveillance in a surveillance-dominated world
2013-11-1_2-04	O'Byrne et al.	Anonymous HIV testing and public health in Ontario, Canada: Understanding HIV surveillance

2013-11-1_2-05	Saltes	'Abnormal' bodies on the borders of inclusion: Biopolitics and the paradox of disability surveillance
2013-11-1_2-06	Fonio et al.	Surveillance, repression and the welfare state: Aspects of continuity and discontinuity in post-Fascist Italy
2013-11-1_2-07	Hintjens	Screening in or out? Selective non-surveillance of unwanted humanity in EU cities
2013-11-1_2-08	Boersma	Liminal surveillance: An ethnographic control room study during a local event
2013-11-1_2-09	Kroener	'Caught on camera': The media representation of video surveillance in relation to the 2005 London Underground Bombings
2013-11-1_2-10	Germain et al.	A prosperous 'business' : The success of CCTV through the eyes of international literature.
2013-11-1_2-11	Gad et al.	A Closed Circuit Technological Vision: On Minority Report, event detection, and enabling technologies
2013-11-1_2-12	Whitson	Gaming the Quantified Self
2013-11-1_2-13	Cakici	Sustainability through surveillance: ICT discourses in design documents
2013-11-1_2-14	Smith et al.	Comparison of survey findings from Canada and the USA on surveillance and privacy from 2006 and 2012
2013-11-1_2-15	Jurgenson	Review of Bauman and Lyon's Liquid Surveillance: A Conversation
2013-11-1_2-16	Grommé	Review of Magnet's When Biometrics Fail: Gender, Race, and the Technology of Identity
2013-11-1_2-17	Saulnier	Review of Gilliom and Monahan's SuperVision: An Introduction to the Surveillance Society
2013-11-1_2-18	Menichelli	Review of Larsen's Setting the Watch: Privacy and the Ethics of CCTV Surveillance
2013-11-3-01	Smith et al.	Surveillance texts and textualism: Truth-telling and trustmaking in an uncertain world
2013-11-3-02	Marks	Monitoring the invisible: Seeing and unseeing in China Mieville's The City & The City
2013-11-3-03	Kruger	Image and exposure: Envisioning genetics as a forensic-surveillance matrix
2013-11-3-04	Edmond et al.	Justicia's Gaze: Surveillance, evidence and the criminal trial
2013-11-3-05	Dee	Welfare Surveillance, Income Management and New Paternalism in Australia
2013-11-3-06	Rao et al.	Subverting ID from above and below: The uncertain shaping of India's new instrument of e-governance
2013-11-3-07	Gregg	Spouse-busting: Intimacy, adultery, and surveillance technology
2013-11-3-08	Westcott et al.	Friendship and trust in the social surveillance network
2013-11-3-09	Negishi	From surveillant text to surveilling device: The face in urban transit spaces
2013-11-3-10	McKay	Covert: the artist as voyeur
2014-11-4-01	Warren et al.	Surveillance and Sport
2014-11-4-02	Manley et al.	Liquid London: Sporting Spectacle, Britishness and Ban-optic Surveillance
2014-11-4-03	Eisenhauer et al.	Fifa-isation: Spatial security, sponsor protection and media management at the 2010 World Cup
2014-11-4-04	Whelan	Surveillance, security and sports mega events: Toward a research agenda on the organisation of security networks
2014-11-4-05	Mason	Watching the watchers: A visual sociology essay on control, security and surveillance at Olympic Women's Soccer in Glasgow
2014-11-4-06	Dixon	The role of surveillance in the construction of authentic football fandom practice.
2014-11-4-07	Warren et al.	Surveillance, governance and professional sport
2014-11-4-08	Jakubowska	Gender verification in sport as a surveillance practice: Inside and outside perception
2014-11-4-09	Rees et al.	The surveillance of racing cyclists in training: A Bourdieusian perspective
2014-11-4-10	Teetzel et al.	From silence to surveillance: Examining the aftermath of a Canadian university doping scandal
2014-11-4-11	Henning	(Self-)Surveillance, anti-doping, and health in non-elite road running
2014-11-4-12	Cook et al.	Monitoring in youth sport: A PARADIGM SHIFT

2014-11-4-13	Murakami Wood	Review of Lefait's Surveillance on Screen
2014-11-4-14	van der Ploeg	Review of Barnard-Wills' Surveillance and Identity
2014-11-4-15	McVey	Review of Andrejevic's Infoglut
2014-12-1-01	Krasmann et al.	'My fingerprint on Osama's cup.' On objectivity and the role of the fictive regarding the acceptance of a biometric technology
2014-12-1-02	Braverman	Governing the wild: Databases, algorithms, and population models as biopolitics
2014-12-1-03	Pieri	Emergent policing practices: Operation Shop a Looter and urban space securitisation in the aftermath of the Manchester 2011 riots
2014-12-1-04	Hallinan et al.	Neurodata and neuroprivacy: Data protection outdated?
2014-12-1-05	Mitchener-Nissen	Failure to collectively assess surveillance-oriented security technologies will inevitably lead to an absolute surveillance society
2014-12-1-06	Clarke	Surveillance by the Australian Media, and Its Regulation
2014-12-1-07	Hardy	Shift recording in residential child care – Purposes, issues and implications for policy and practice
2014-12-1-08	Myers et al.	Leprosy and the Plague: State Surveillance of Low-Income Fathers
2014-12-1-09	Huey	The Problem with Ethics: Difficulties in Constructing Normative Frameworks
2014-12-1-10	Macnish	Just Surveillance? Towards a Normative Theory of Surveillance
2014-12-1-11	Hosein	On Just Surveillance
2014-12-1-12	Stoddart	Challenging 'Just Surveillance Theory': A Response to Kevin Macnish's 'Just Surveillance? Towards a Normative Theory of Surveillance'.
2014-12-1-13	Palm	Conditions under which Surveillance may be Ethically Justifiable: Remarks on Kevin Macnish's proposed normative theory of surveillance
2014-12-1-14	Marx	Toward an Imperial System of Surveillance Ethics
2014-12-1-15	Macnish	Response
2014-12-1-16	Chadderton	Review of Taylor's Surveillance Schools
2014-12-2-01	Andrejevic et al.	Big Data Surveillance: Introduction
2014-12-2-02	van Dijck	Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology
2014-12-2-03	Degli Esposti	When big data meets dataveillance: the hidden side of analytics
2014-12-2-04	French	Gaps in the gaze: Informatic practice and the work of public health surveillance
2014-12-2-05	Reigeluth	Why data is not enough: Digital traces as control of self and self-control
2014-12-2-06	van Otterlo	Automated experimentation in Walden 3.0: The next step in profiling, predicting, control and surveillance
2014-12-2-07	Klauser	From self-tracking to smart urban infrastructures: Towards an interdisciplinary research agenda on Big Data
2014-12-2-08	Thomas	Pandemics of the future: Disease surveillance in real time
2014-12-2-09	Boersma	Review of Trottier's Social Media as Surveillance: Rethinking Visibility in a Converging World
2014-12-2-10	Fang	Review of Rosen and Santesso's The Watchman in Pieces: Surveillance, Literature, and Liberal Personhood
2014-12-2-11	Galdon Clavell	Review of Bakir's Torture, Intelligence and Sousveillance in the War on Terror
2014-12-3-01	Whitson et al.	Game Studies meets Surveillance Studies at the Edge of Digital Culture: An Introduction to a special issue on Surveillance, Games and Play
2014-12-3-02	Kerr et al.	Surveillant Assemblages of Governance in Massively Multiplayer Online Games: A Comparative Analysis
2014-12-3-03	Collister	Surveillance and Community: Language Policing and Empowerment in a World of Warcraft Guild
2014-12-3-04	O'Donnell	Getting Played: Gamification, Bullshit, and the Rise of Algorithmic Surveillance
2014-12-3-05	Andersen	Games of Drones: The Uneasy Future of the Soldier-Hero in Call of Duty: Black Ops II
2014-12-3-06	Farman	Creative misuse as resistance: Surveillance, mobile technologies, and locative games
2014-12-3-07	Hulsey et al.	The gift that keeps on giving: Google, Ingress, and the gift of surveillance

2014-12-3-08	Barron	'I had no credit to ring you back': Children's strategies of negotiation and resistance to parental surveillance via mobile phones
2014-12-3-09	Rayness-Goldie et al.	Gaming privacy: A Canadian case study of a co-created privacy literacy game for children
2014-12-3-10	Cybulski	Enclosures at Play: Surveillance in the Code and Culture of Videogames
2014-12-3-11	Canossa	Reporting from the snooping trenches: Changes in attitudes and perceptions towards behavior tracking in digital games
2014-12-3-12	Walker	Watching us play: Postures and platforms of live streaming
2014-12-3-13	Gangneux	Diverted and diverting glances at cameras: Playful and tactical approaches to surveillance
2014-12-3-14	Robbins et al.	Playdates with Big Brother: Playfully repurposing surveillance cameras to build communities
2014-12-3-16	Frois	Review of Björklund and Svenonius' Video Surveillance and Social Control in a Comparative Perspective
2014-12-3-17	Staples	Review of Lianos' The New Social Control: The Institutional Web, Normativity and the Social Bond
2014-12-4-01	Hasselberg	Coerced to leave: Punishment and the surveillance of foreign-national offenders in the UK
2014-12-4-02	Monaghan	Security traps and discourses of radicalization: Examining surveillance practices targeting Muslims in Canada
2014-12-4-03	Schaefer et al.	Watching the watchers and McLuhan's Tetrad: The limits of cop-watching in the internet age
2014-12-4-04	Hristova	Recognizing friend and foe: Biometrics, veridiction, and the Iraq War
2014-12-4-05	Rodriguez Tejada	Surveillance and student dissent: The case of the Franco dictatorship
2014-12-4-06	Thompson	Making up soldiers: The role of statistical oversight and reactive path dependence in the effectiveness of Canada's WWII Mobilization Program 1940-1943
2014-12-4-07	Fisk	"...when no one is hearing them swear" - Youth safety and the pedagogy of surveillance
2014-12-4-08	Oostveen et al.	Child location tracking in the US and the UK: Same technology, different social implications
2014-12-4-09	Morgan	Surveillance in contemporary health and social care: friend or foe?
2014-12-4-10	Barnard-Wills	Review of Ajana's Governing through Biometrics
2014-12-4-11	Clément	Review of Larsen and Walby's 'Brokering Access'
2014-12-4-12	Trottier	Review of Leistert's From Protest to Surveillance - The Political Rationality of Mobile Media
2014-12-4-13	Klauser	Review of Zurawski's Raum – Weltbild – Kontrolle. Raumvorstellungen als Grundlagesgesellschaftlicher Ordnung und ihrer Überwachung
2015-13-1-01	Singh et al.	Doing Surveillance Studies (Part II): Critical approaches to methodology and pedagogy
2015-13-1-02	McCoy	Policing the imperial periphery: The Philippine-American War and the origins of U.S. global surveillance
2015-13-1-03	Green et al.	Surveillance and ethnography: Researching surveillance as everyday life
2015-13-1-04	Abu-Laban	Gendering Surveillance Studies: The empirical and normative promise of feminist methodology
2015-13-1-05	Phillips	Work and play at the threshold of legibility: Theatre as method and pedagogy in surveillance research
2015-13-1-06	Arteaga	Doing Surveillance Studies in Latin America: The insecurity context
2015-13-1-07	Cohen	Studying Law studying surveillance
2015-13-1-08	Stoddart	The Named Person: Surveillance and the wellbeing of children and young people in Scotland
2015-13-1-09	Mehrabov	Exploring Terra Incognita: Mapping surveillance studies from the perspective of media and communication research
2015-13-1-10	Spiller	Review of Friedewald and Pohoryles' Privacy and Security in the Digital Age
2015-13-1-11	McCahill	Review of Smith's Opening the Black Box: The Work of Watching

2015-13-2-01	Murakami Wood et al.	Before and after Snowden
2015-13-2-02	Lyon	The Snowden stakes: Challenges for understanding surveillance today
2015-13-2-03	Verde Garrido	Contesting a biopolitics of information and communications: The Importance of truth and sousveillance after Snowden
2015-13-2-04	Keiber	Surveillance hegemony
2015-13-2-05	van der Velden	Leaky apps and data shots: Technologies of leakage and insertion in NSA-surveillance
2015-13-2-06	Schulze	Patterns of surveillance legitimization: The German discourse on the NSA scandal
2015-13-2-07	Evans	'The footage is decisive': Applying the thinking of Marshall McLuhan to CCTV and police misconduct
2015-13-2-08	Greene	Drone vision
2015-13-2-09	Johnson	Surveillance, Pastoral power and embodied infrastructures of care among migrant Filipino Muslims in the Kingdom of Saudi Arabia
2015-13-2-10	Wright et al.	"It depends on who you are, what you are": 'Community Safety' and sex workers' experience with surveillance
2015-13-2-11	Vredenburg	Notes toward a meteorology of the cloud
2015-13-2-12	Lippert	Thinking about law and surveillance
2015-13-2-13	Austin	Surveillance and the rule of law
2015-13-2-14	Warren	Surveillance, criminal law and sovereignty
2015-13-2-15	Williams	Law, surveillance, and financial markets
2015-13-2-16	Braverman	Hyperlegality and heightened surveillance: The case of threatened species lists
2015-13-2-17	Fitzpatrick et al.	Foucault, surveillance and the Law of the Outside
2015-13-2-18	Menichelli	Review of Kitchin's 'The Data Revolution'
2015-13-2-19	Nellis	Review of Ridker's 'Privacy Policy'

Appendix B: Articles excluded during the S&S Corpus compilation

During the data collection four minor issues complicated the access to the correct files and I resolved them as follows:

- At the time of the corpus compilation, the PDF file for article 2010-07-2-09 was missing from the website, although the article was listed there. The file became available later and was copy-pasted into a text file and manually cleaned.
- Two articles listed on the website were linked to wrong files (of previous articles from the journal): 2005-03-1-05 (duplicate of 2005-02-1-05) and 2006-04-1_2-01 (duplicate of 2006-04-1_2-02). The duplicate files were discovered via concordance lines and removed. PDF files of the original articles were accessed from the journal's old website (<http://www.surveillance-and-society.org/journal.htm>), and, as in (i), copy-pasted and manually cleaned. Two obvious errors of merged words (triggered by copy-pasting) were corrected manually during this process.¹
- Two separate book reviews were published twice with no changes to the contents, but minor changes in layout and the reference of the reviewed book. The two duplicate files were removed: 2009-06-1-12 (duplicate of 2008-05-3-08) and 2009-06-4-11 (duplicate of 2009-06-3-16)

¹ The first merged instance Itwasalsoknownasthe“Indianlist”.Touse theominouswordsofEdwinBlack(2001: 92) was corrected to It was also known as the “Indian list”. To use the ominous words of Edwin Black (2001: 92). The second merged instance “Non-confidencewasgeneratedbycategoricalcomparisonandcollapse ratherthan” was corrected to “Non-confidence was generated by categorical comparison and collapse rather than”

Appendix C: DEFINE concordance of *surveillance*

Concordance 1: All 124 instances of surveillance co-occurring (span of five words) with *define/defines/defined/defining/definition* (excluding overlapping and irrelevant lines)

1	better definition of the new	surveillance	is the use of technical mean	2002-01-1-02
2	. This definition of the new	surveillance	excludes the routine, non-te	2002-01-1-02
3	ot without struggle. Even if	surveillance	by definition always involve	2002-01-1-02
4	the dictionary definition of	surveillance	as "close observation, espec	2002-01-1-02
5	he Concise Oxford Dictionary	surveillance	is defined as "close observa	2002-01-1-02
7	ex in its claims. He defines	surveillance	as the collection and storag	2003-01-2-02
8	form of maintaining control.	Surveillance	- by definition - involves t	2003-01-3-05
9	s. First, facial recognition	surveillance	is defined, highlighting the	2003-01-3-06
10	n fit our definition of what	surveillance	is? In my opinion, for now,	2003-01-4-04
11	s. He states: "A relation of	surveillance	, defined and regulated, is i	2003-01-4-05
14	cient to define it as camera	surveillance	of publicly accessible space	2004-02-2_3-07
16	e disciplinary society, such	surveillance	captures and defines the sub	2004-02-2_3-19
17	the view that the complex of	surveillance	, as defined above, serves as	2004-02-2_3-19
18	Surveillance - a definition	surveillance	can take many forms from the	2005-03-1-07
20	ontrast between 'traditional	surveillance	,' defined as "close observat	2005-03-2_3-02
21	Marx (2003: 370), we define	surveillance	as those technologies of vis	2005-03-2_3-02
22	reed upon definition of what	surveillance	is. Marx's (2005, 2001) iter	2005-03-2_3-03
23	of the defining elements in	surveillance	, modern or otherwise, but th	2005-03-2_3-08
24	e modified the definition of	surveillance	offered above, to 'the purpo	2005-03-2_3-10
28	Maguire have defined police	surveillance	as 'a wide range of methods	2007-04-3-05
29	cause. The law defined when	surveillance	was reasonable and provided	2008-05-1-10
31	the definition and scope of	Surveillance	Studies, looks at surveillan	2009-06-1-08
32	been a defining attribute of	surveillance	medicine; now such self exam	2009-06-2-01
33	definition of public health	surveillance	and its continued relevance	2009-06-2-01
34	two recent documents define	surveillance	in a public health context i	2009-06-2-02
35	rkelman define public health	surveillance	, these contributors, whoever	2009-06-2-02
36	luding, in his definition of	surveillance	, the idea that it should inv	2009-06-2-02
37	following definition: Health	Surveillance	may be defined as the tracki	2009-06-2-02
38	[...] Langmuir defined disease	surveillance	as 'the continued watchfulne	2009-06-2-02
40	circumscribed definition of	surveillance	, which I will discuss below.	2009-06-2-02
41	down the basic definition of	surveillance	in a public health context.	2009-06-2-02
42	ance studies scholars define	surveillance	differently than the way it	2009-06-2-02
43	overnance. In public health,	surveillance	, as defined by Alexander Lan	2009-06-2-03
44	eal time, but via prosthetic	surveillance	regulates and defines bodies	2009-06-2-07
45	of what has been defined as	'surveillance	medicine': This new surveill	2009-06-2-07
46	endency towards a prosthetic	surveillance	, which regulates and defines	2009-06-2-07
47	w from numerous studies that	surveillance	, as defined above, can origi	2009-06-3-01
48	al surveillance Lyon defines	surveillance	as "any collection and proce	2009-06-3-03
50	chnologies and practices of	surveillance	helped define the national a	2009-07-1-01
51	at surveillance. Therefore,	surveillance	will be defined more broadly	2009-07-1-03
53	ve to define 'the nature' of	surveillance	in general terms. Instead, i	2009-07-1-05
54	icult to define. Apparently,	surveillance	is at work in every corner o	2009-07-1-05
55	a to define surveillance qua	surveillance	, conceptual confusion will c	2009-07-1-07
57	n surveillance must show how	surveillance	possesses defining features	2009-07-1-07
58	efine the next generation of	Surveillance	Studies, and I offer encoura	2009-07-1-08
59	enting something defined as	'Surveillance	Studies' - matters to the fu	2009-07-1-08
60	presentative forms where the	surveillance	definition is clear and invo	2010-07-3_4-02
61	context The legal context of	surveillance	is defined by the United Sta	2010-08-1-02
63	ation defined by a matrix of	surveillance	that includes border patrols	2010-08-2-02
64	led through data collection.	Surveillance	has been defined as: "any co	2010-08-2-03
65	hat expand the definition of	surveillance	. Technology movements, vario	2010-08-2-03
66	to the question of defining	surveillance	which allow for the manifold	2010-08-2-09
67	in particular, how to define	surveillance	. In this article, we do not	2010-08-2-09
68	e problem of definition. The	surveillance	literature offers many defin	2010-08-2-09
69	the most cited definition of	surveillance	proposed by Lyon (2001) suff	2011-08-3-14
70	the nature and definition of	surveillance	. When reading one is constan	2011-08-3-14
71	y also demonstrates how this	surveillance	uses race to define the bord	2011-08-4-02
72	gue that practices of direct	surveillance	define the border between th	2011-08-4-02
73	by mediated visibilities and	surveillance	, and it is not necessary to	2011-08-4-03
74	the defining facets of both	surveillance	practice and modern subjecti	2011-08-4-03
75	in the definition of 'covert	surveillance	', RIPA provides a regulatory	2011-08-4-05
76	interpreted broadly. Covert	surveillance	is defined as that carried o	2011-08-4-05
77	station defines as 'directed	surveillance	'. RIPA states that surveilla	2011-08-4-05
80	of surveillance. They define	surveillance	as involving 'the collection	2011-08-4-06

81	rting point for showing that	surveillance	and audit can be defined as	2011-08-4-06
82	a rather wide definition of	surveillance.	They define surveillance as	2011-08-4-06
83	is best defined in terms of	surveillance,	the social and individual v	2011-08-4-08
84	notes the modern landscape.	Surveillance	as a definition of, or frame	2011-08-4-08
85	ould be defined in terms of	surveillance	because it more accurately f	2011-08-4-12
86	eving its primary goal, i.e.	surveillance	as it has been defined. Last	2011-09-1_2-07
87	e spend less effort defining	surveillance	studies, as this literature	2012-09-3-05
88	the intensification of urban	surveillance	may be defined less by a pri	2012-09-3-05
90	the intensification of urban	surveillance	therefore may be less a priv	2012-09-3-05
91	f Lyon's definition reveals,	surveillance	has the aim of influencing o	2012-09-4-03
92	David Lyon (2001, 2) defines	surveillance	as "the collection and proce	2012-09-4-03
93	to Lyon's definition, social	surveillance	certainly involves "the focu	2012-09-4-04
94	Overview, David Lyon defines	surveillance	as "the focused, systematic	2012-09-4-04
95	o surveillance. Lyon defines	surveillance	as "any collection and proce	2012-09-4-06
97	oal of defining the place of	surveillance	in a region, further than a	2012-10-1-01
98	ernization of the region. If	surveillance	in modern societies is defin	2012-10-1-01
100	lance agenda and in defining	surveillance	priorities and risks, but th	2012-10-1-10
104	itly inclusive definition of	surveillance	as encompassing "all forms o	2012-10-2-02
105	pite the broad definition of	surveillance	as involving "the collection	2012-10-2-02
108	defines as 'social sorting,'	surveillance	is never neutral, since it e	2012-10-3_4-09
109	ty and Ericson (2000, 2006),	surveillance	has been defined as the '...	2012-10-3_4-09
110	surveillance in daily life,	surveillance	stands out as a defining fea	2012-10-3_4-09
112	se Dictionary, 2013, defines	surveillance	as follows: sur·veil·lance [2013-11-1_2-03
113	. Bearing this definition of	surveillance	and the work of Hardt and Ne	2013-11-1_2-04
114	n's (2007) definition: while	surveillance	'is the focused, systematic	2013-11-1_2-04
115	(2007), I define disability	surveillance	as the practice of collectin	2013-11-1_2-05
118	idely accepted definition of	surveillance	proposed by Lyon (2007), I d	2013-11-1_2-05
120	rotection, we are looking at	surveillance."	(2006: 4) This definition	2013-11-1_2-07
121	6: 6). A broad definition of	surveillance	from the 2006 Surveillance S	2013-11-1_2-07
122	2011). 2) Normative: liminal	surveillance	defines what can be seen as	2013-11-1_2-08
123	of policing. By definition,	Surveillance	Studies is concerned with al	2013-11-1_2-10
124	ce of categories by defining	surveillance	itself as a process of trans	2013-11-1_2-13
125	ing David Lyon's definition,	surveillance	takes the form of routine at	2013-11-1_2-13
127	uthors note that they define	surveillance	broadly as "monitoring peopl	2013-11-1_2-17
128	nature' of the definition of	surveillance	based on the idea of 'close	2013-11-3-02
129	aracteristic of contemporary	surveillance	society (Lyon 2010), or even	2013-11-3-04
130	able to define the powers of	surveillance	or, indeed, to devise a mean	2013-11-3-05
132	ple, Lyon (2007: 14) defines	surveillance	as 'the focused, systematic	2014-11-4-04
133	ermanent' (Zedner 2009: 19).	Surveillance	can be defined in a similar	2014-11-4-04
135	, it has been suggested that	surveillance,	defined as 'watching over'	2014-11-4-12
136	established ways of defining	surveillance	also suggest the need to re-	2014-12-2-01
137	i Wood and colleagues define	surveillance	as, "purposeful, routine, sy	2014-12-2-01
138	tion of data-mining enhanced	surveillance.	What is significant about t	2014-12-2-01
139	ces, including public health	surveillance,	which has been defined as "	2014-12-2-04
140	humans As mentioned before,	surveillance	is commonly defined as an 'e	2014-12-2-07
141	in current Big Data trends.	Surveillance	is commonly defined as the '	2014-12-2-07
143	l shift in the definition of	surveillance	in a public health context t	2014-12-2-08
144	st many others, reminds us,	Surveillance	can be defined as the system	2014-12-3-01
146	participatory nature of the	surveillance	was a defining aspect of the	2014-12-3-03
147	the concept of participatory	surveillance,	defined using a combination	2014-12-3-03
148	d English Dictionary defines	surveillance	as: 'Watch or guard kept ove	2014-12-4-09
149	o secure any definition for	surveillance'	of course exceeds the scope	2015-13-1-03
150	ncy is that the relations of	surveillance	under scrutiny again become	2015-13-1-03
151	h fields or sites defined as	surveillance	in an a priori way (via the	2015-13-1-03
152	hips in settings defined as	'surveillance'	in an a priori fashion. The	2015-13-1-03
153	hips in settings defined as	'surveillance'	in an a priori fashion. As	2015-13-1-03
154	es that does not investigate	surveillance	defined as a set of 'things'	2015-13-1-03
157	ltitude. By definition, mass	surveillance	means that anyone and everyo	2015-13-2-02
158	urveillance, Migration, Care	Surveillance	is defined generally as a se	2015-13-2-09
160	has employed a definition of	surveillance	that encompasses "all forms	2015-13-2-16

Concordance 2: All 13 instances of overlapping DEFINE concordance lines (excluded from analysis)

6	cteristic of much of the new	surveillance.	A better definition of the	2002-01-1-02
19	Ramia, 2002; Gilliom, 2001).	Surveillance	- a definition Surveillance	2005-03-1-07
39	ice it to say here that most	surveillance	studies scholars define surv	2009-06-2-02
49	xt. Theorizing technological	surveillance	Lyon defines surveillance as	2009-06-3-03
52	ebly complicating attempts at	surveillance.	Therefore, surveillance wil	2009-07-1-03
56	ar set of criteria to define	surveillance	qua surveillance, conceptual	2009-07-1-07
79	veillance. Agents, Audit and	Surveillance	Defining a relationship as i	2011-08-4-06
96	rveillance and resistance to	surveillance.	Lyon defines surveillance a	2012-09-4-06

101	en successful in shaping the	surveillance	agenda and in defining surve	2012-10-1-10
103	eady entangled in systems of	surveillance .	Broadening the definition o	2012-10-2-01
106	captive animals. Redefining	Surveillance	Despite the broad definition	2012-10-2-02
126	ity is linked to ICT through	surveillance .	Following David Lyon's defi	2013-11-1-2-13
159	grant and diasporic context.	Surveillance ,	Migration, Care Surveillanc	2015-13-2-09

Concordance 3: All 23 instances of “irrelevant” concordance lines (excluded from analysis)

12	ion of a 'closed circuit' of	surveillance	of a defined population. Int	2004-02-1-01
13	ined minimum. This selective	surveillance	of certain 'spatial points'	2004-02-2_3-03
15	dimension of the space under	surveillance .	Potential clients are there	2004-02-2_3-18
25	s for exploring the power of	surveillance	technologies to define what	2006-04-1_2-06
26	esponse network' surrounding	surveillance	cameras. This network is def	2007-04-3-04
27	where defined. Consequently	'surveillance	creep' has gone undetected a	2007-04-3-05
30	an even ask why a state uses	surveillance	mechanisms, we need to defin	2009-06-1-05
62	ure guide both groups use of	surveillance	and definition of migration	2010-08-2-02
78	business actors as agents of	surveillance .	Agents, Audit and Surveilla	2011-08-4-06
89	rity and internally directed	surveillance	too. What defines the intens	2012-09-3-05
99	gers that different forms of	surveillance	imply, without defined crite	2012-10-1-02
102	ition of violence to include	surveillance	It is a difficult task to cr	2012-10-2-01
107	the participatory nature of	surveillance .	He defines the project of h	2012-10-2-03
111	ned above: people understand	surveillance	(they assign meanings to it)	2013-11-1_2-02
116	g a discussion of disability	surveillance	is to provide a definition o	2013-11-1_2-05
117	on of disability, disability	surveillance	involves monitoring bodies a	2013-11-1_2-05
119	imply defined as 'unwanted'.	Surveillance	for unwanted people, of whom	2013-11-1_2-07
131	e definition exists) through	surveillance	and other strategies of enfo	2013-11-3-05
134	ionship between security and	surveillance .	Security has traditionally	2014-11-4-04
142	simultaneous. In its disease	surveillance	documents, the WHO defines r	2014-12-2-08
145	principle of chess - unlike	surveillance	- is that both players have	2014-12-3-01
155	ontemporary Mexico. Criminal	surveillance	The production of drugs from	2015-13-1-06
156	is patently urgent. As with	surveillance	or privacy, defining securit	2015-13-2-02

Concordance 4: All 13 instances of (*surveillance, definitions*) within a span of five words

1	oners, 2002). Definitions of	surveillance	include the 'close observati	2005-03-2_3-10
2	of acceptable behaviour. The	surveillance	of shopping centre users and	2006-04-1_2-03
3	e problem of definition. The	surveillance	literature offers many defin	2010-08-2-09
4	d narrowly cast the problem.	Surveillance	definitions elicit broader s	2011-08-4-08
5	management as an instance of	surveillance .	Current definitions of surv	2012-10-2-02
6	ance. Current definitions of	surveillance	do not suffice to reflect th	2012-10-2-02
7	initions and/or meanings of	surveillance ,	but rather sought (a) to de	2013-11-1_2-04
8	e many definitions of online	surveillance ,	for clarity we situate how	2013-11-3-08
9	conventional definitions of	surveillance	emphasize its systematic and	2014-12-2-01
10	t influential definitions of	surveillance .	In their report on "The Sur	2014-12-2-01
11	wayside, and definitions of	surveillance	that account for such practi	2014-12-2-01
12	and a priori definitions of	surveillance	relations, in a relatively u	2015-13-1-03
13	those specified in standard	surveillance	definitions' (p. 6). Watchin	2015-13-1-11

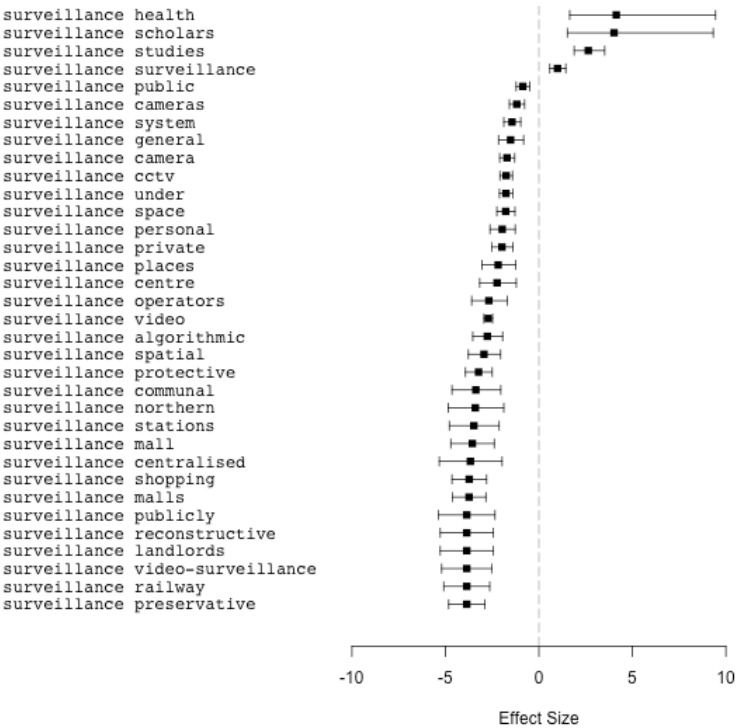
Appendix D: List of issue allocation in Mehrabov's (2015) content analysis*

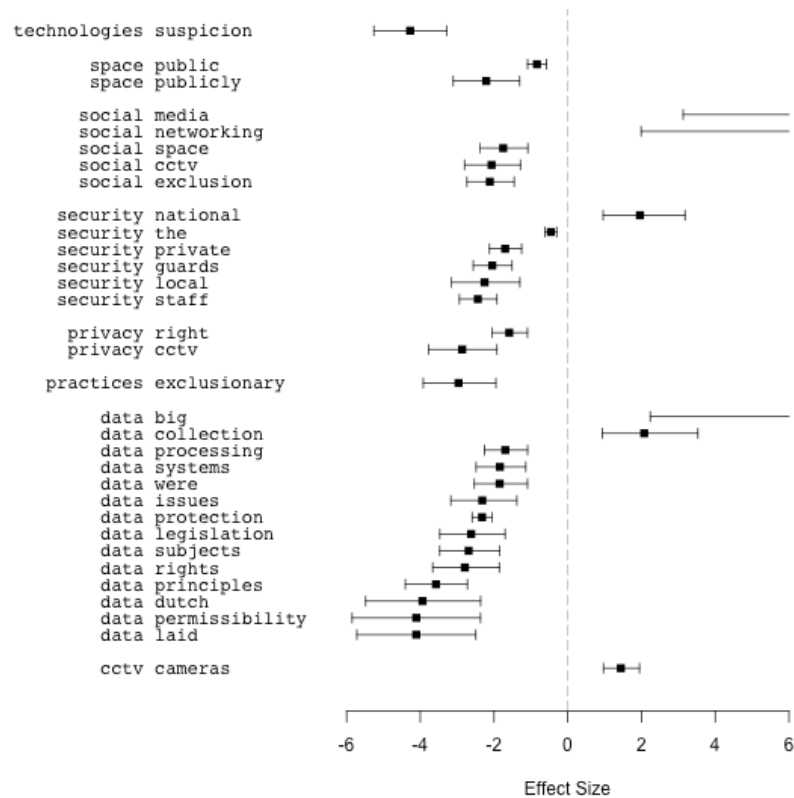
Mehrabov's (2015) theme	Issue no.	Issue title
1. Classic Surveillance: "Discipline" and "Control"	1(3)	Foucault and Panopticism Revisited
	2(2, 3)	The Politics of CCTV in Europe and Beyond
	6(1)	<u>Relaunch Issue</u> : Revisiting Video Surveillance
2. Identity-based Surveillance	4(3)	Surveillance and Criminal Justice: Part 1
	4(4)	Surveillance and Criminal Justice: Part 2
	5(3)	Surveillance and Inequality
	6(2)	Health, Medicine and Surveillance
	6(4)	Gender, Sexuality and Surveillance
	7(3, 4)	Surveillance, Children and Childhood
	9(4)	Cyber-Surveillance in Everyday Life
3. Mobility and Stasis	1(4)	Surveillance and Mobilities
	5(2)	Smart Borders and Mobilities: Spaces, Zones, Enclosures
	9(3)	Urban Surveillance
4. Work, Power and Resistance	1(2)	Work
	2(4)	People Watching People
	6(3)	Surveillance and Resistance
	8(2)	Surveillance and Empowerment
	8(3)	Marketing, Consumption and Surveillance

*Adapted from Mehrabov (2015, pp. 119–122)

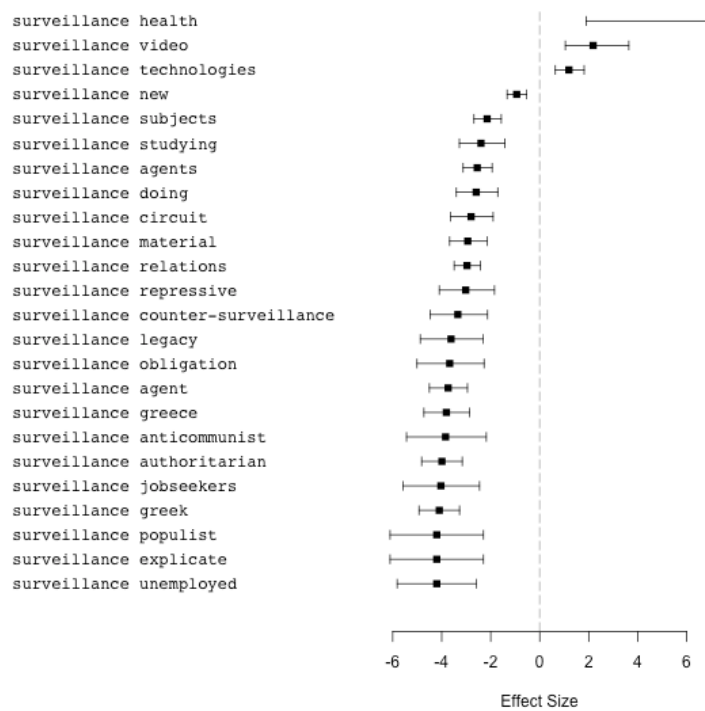
Appendix E: Volume-based co-occurrence results for Chapter 4

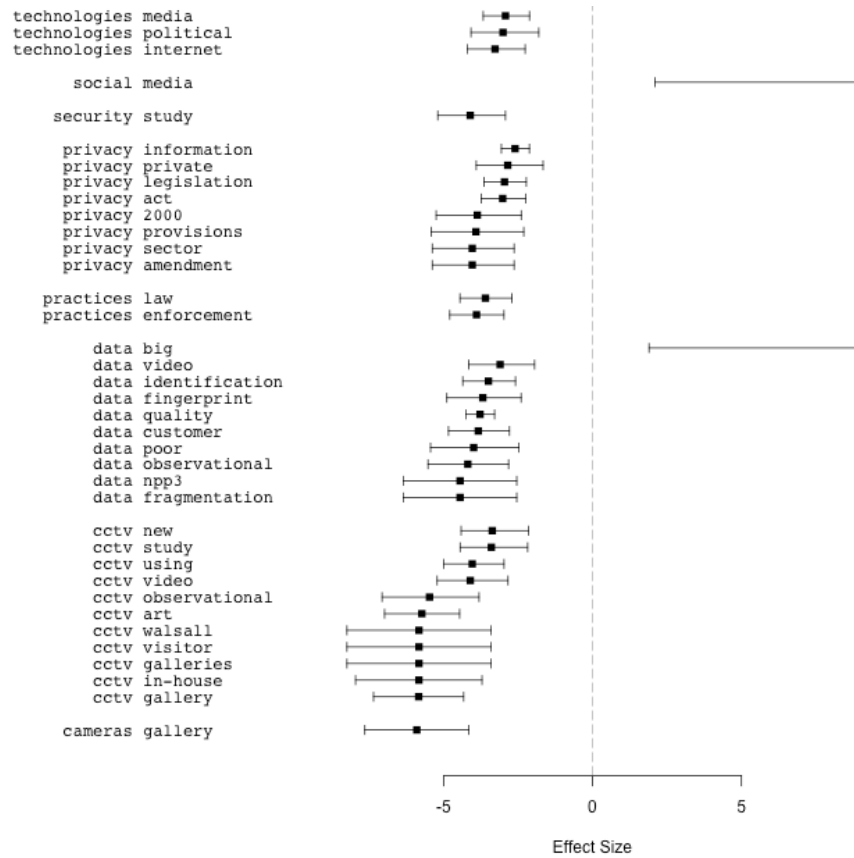
Volume 2 vs the whole S&S Corpus



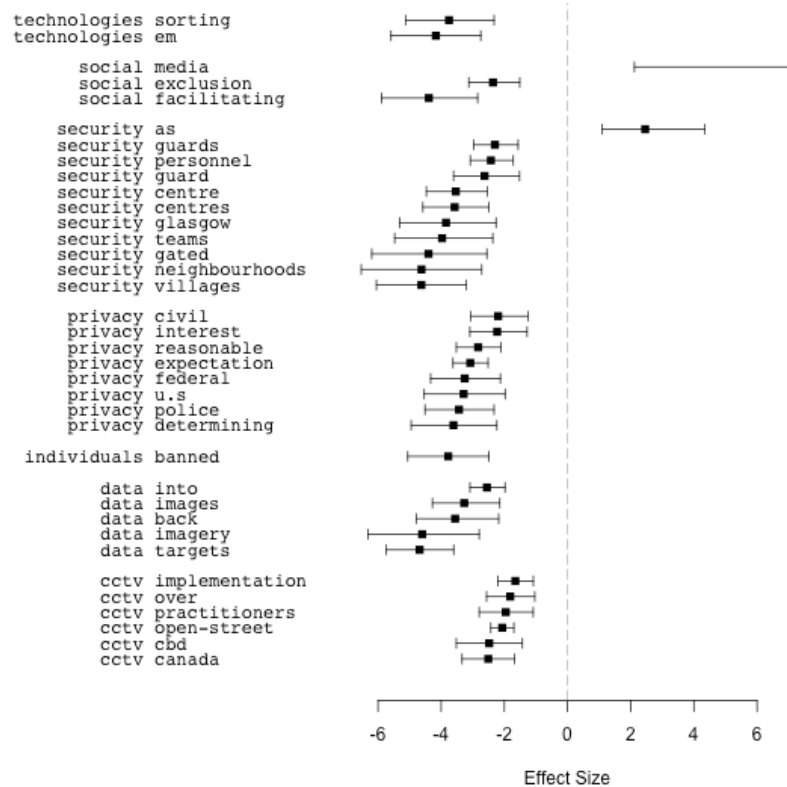
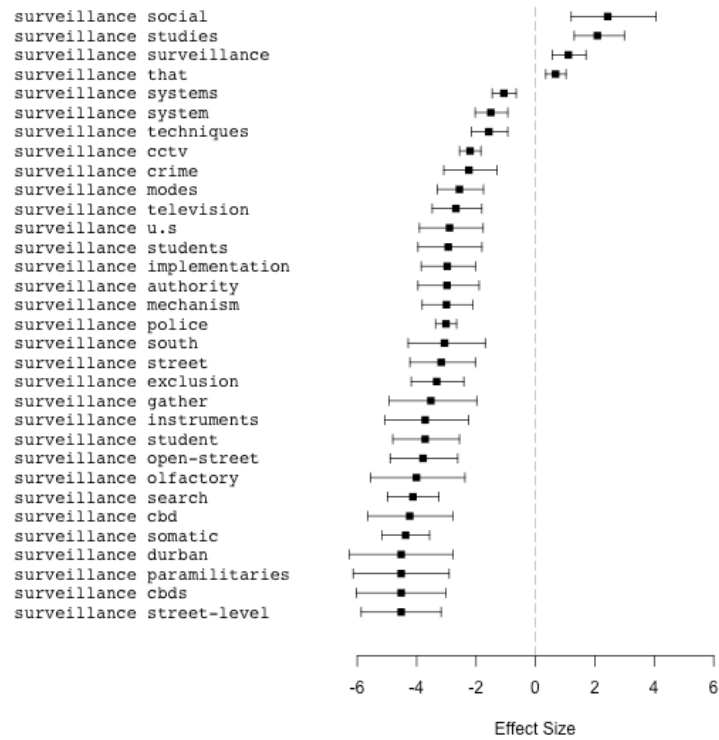


Volume 3 vs. the whole S&S Corpus

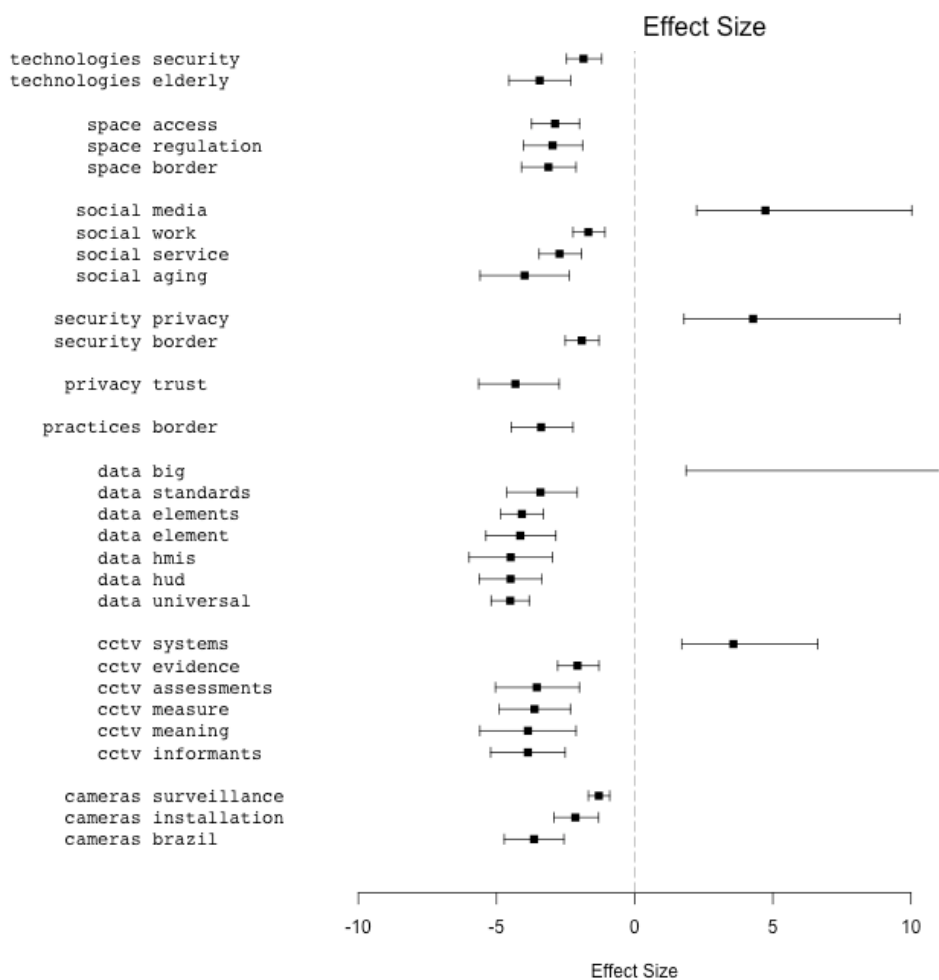
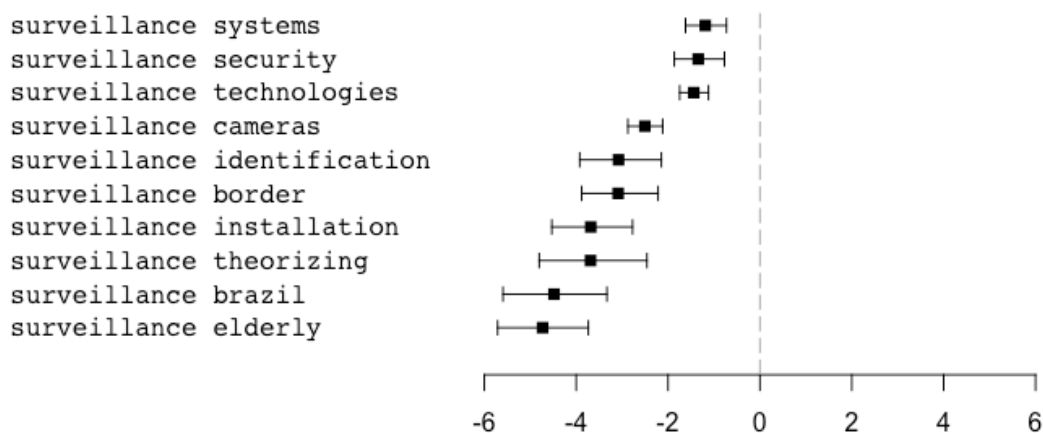




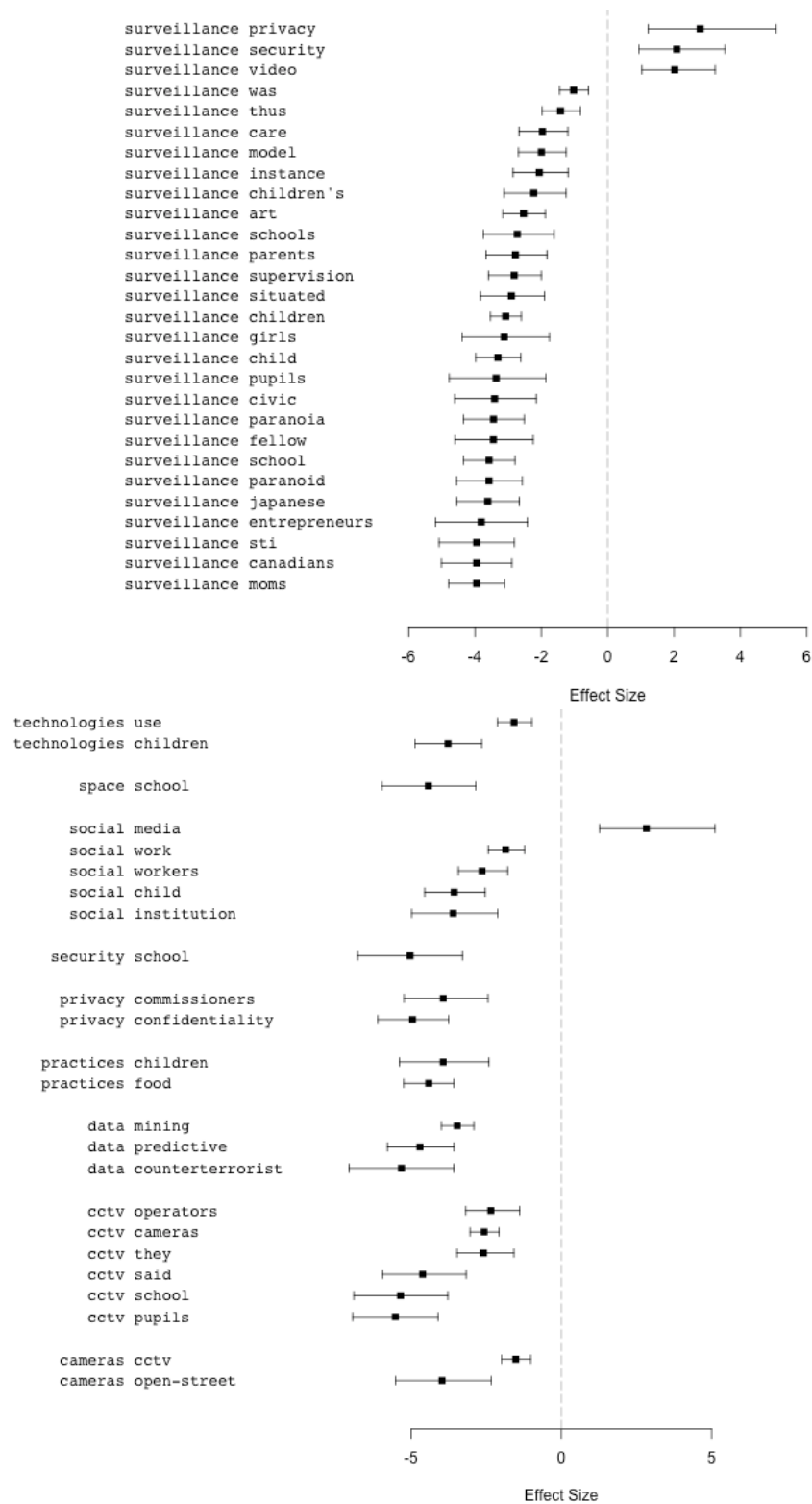
Volume 4 vs. the whole S&S Corpus



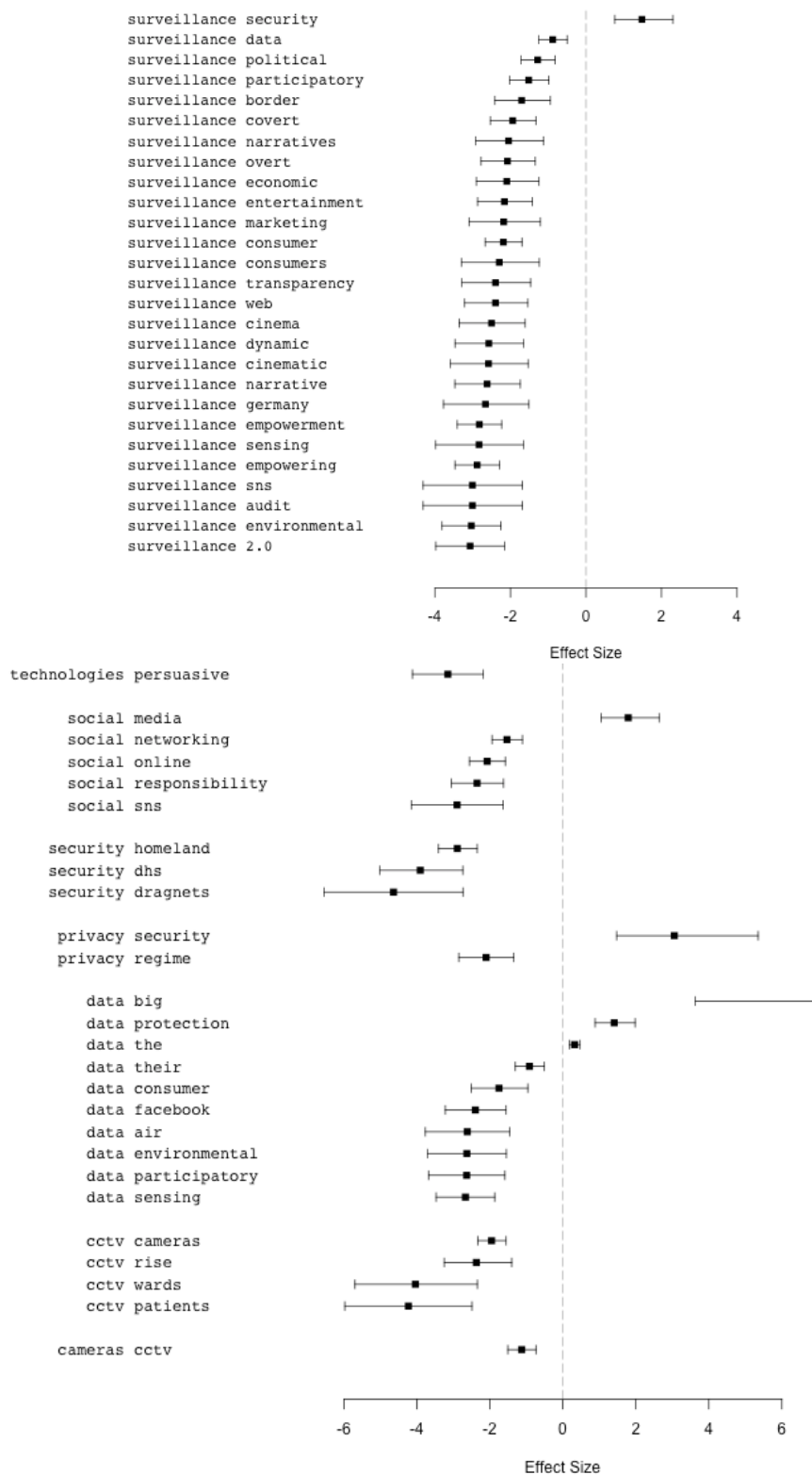
Volume 5 vs. the whole S&S Corpus



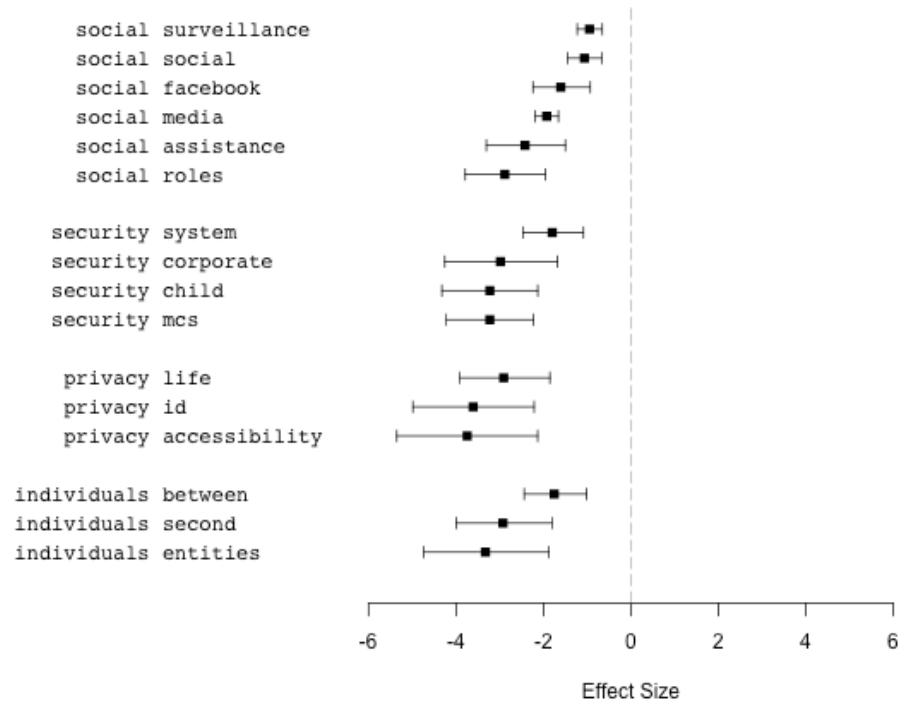
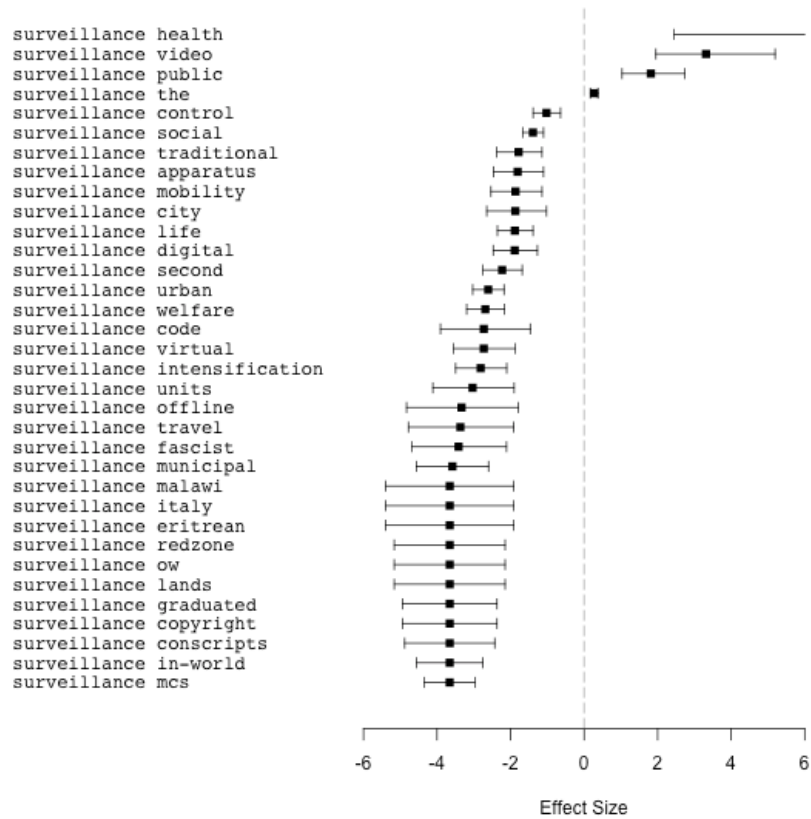
Volume 7 vs. the whole S&S Corpus



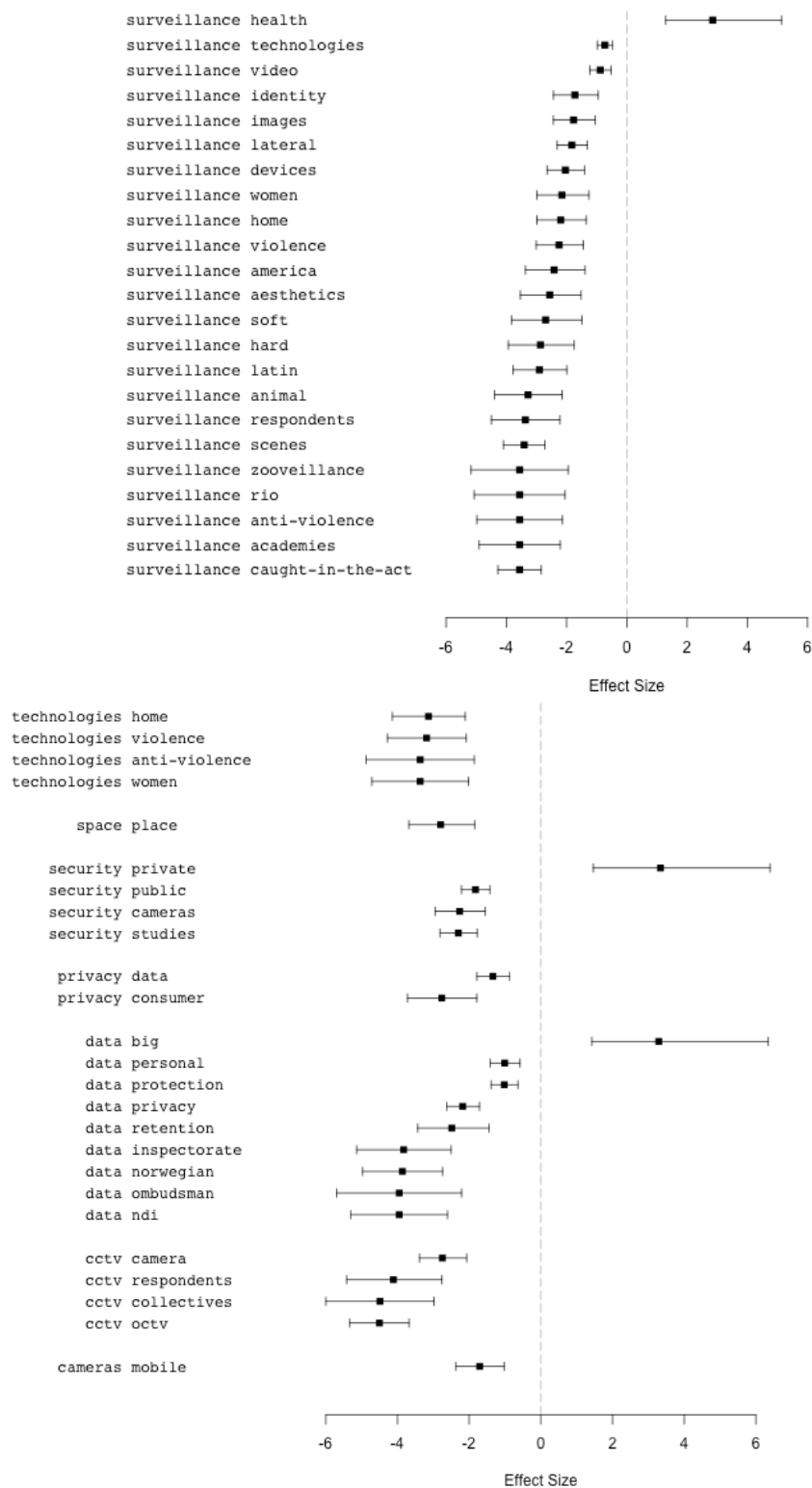
Volume 8 vs. the whole S&S Corpus



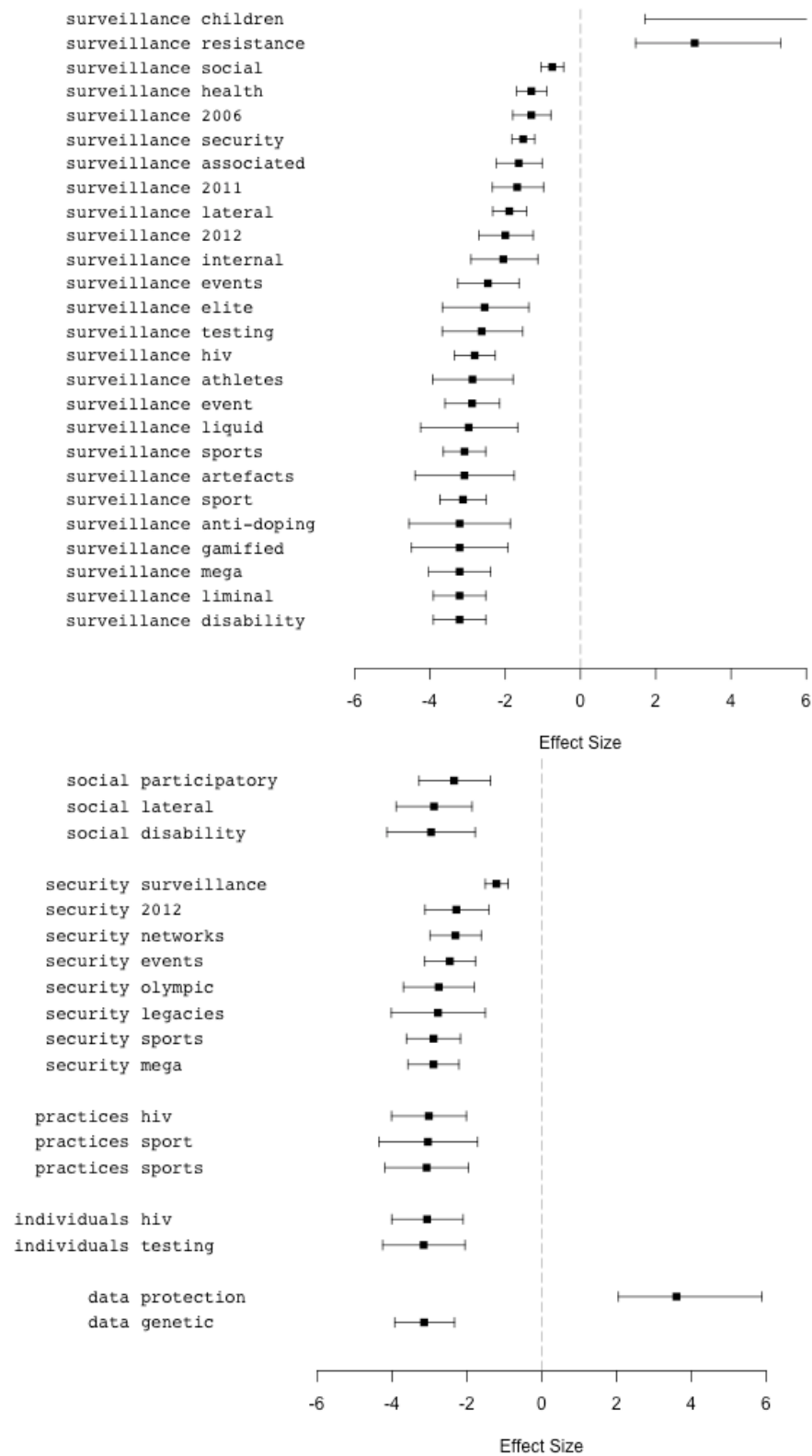
Volume 9 vs. the whole S&S Corpus



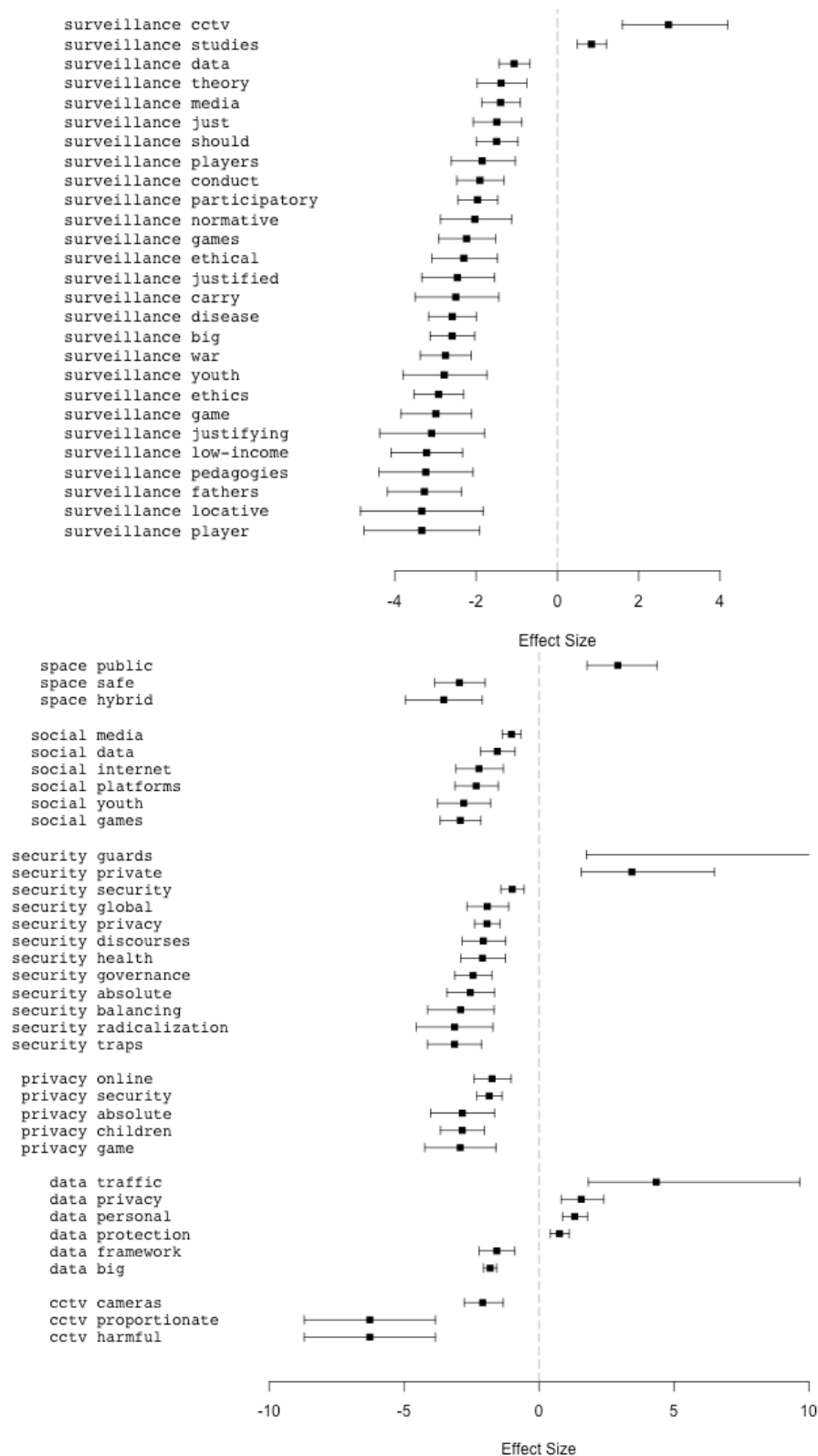
Volume 10 vs. the whole S&S Corpus



Volume 11 vs. the whole S&S Corpus



Volume 12 vs. the whole S&S Corpus



Appendix F: List of URLs in the Surveillance Blog Corpus

File ID	URL
blogspot_outfile000	http://150patrick44.blogspot.com
blogspot_outfile002	http://2013infoethics.blogspot.com
blogspot_outfile003	http://2020cctv.blogspot.com
blogspot_outfile007	http://9-11themotherofallblackoperations.blogspot.com/2014/06/nsa-satellite-signals-intelligence.html
blogspot_outfile008	http://9-11themotherofallblackoperations.blogspot.com/2014/08/the-state-of-new-hampshire-passes-law.html
blogspot_outfile009	http://9507515.blogspot.com
blogspot_outfile010	http://aadhaar-articles.blogspot.com/2012_07_01_archive.html
blogspot_outfile011	http://aadhaar-articles.blogspot.com/2016_03_01_archive.html
blogspot_outfile015	http://advogvr.blogspot.com
blogspot_outfile016	http://akamyp.blogspot.com
blogspot_outfile018	http://alanpetersnewsbriefs.blogspot.com/2007/09/cra-terror-assessment.html
blogspot_outfile019	http://alfa-in-security.blogspot.com
blogspot_outfile021	http://allenshortnacy.blogspot.com
blogspot_outfile023	http://allsecuritysystems.blogspot.com
blogspot_outfile026	http://antifascist-calling.blogspot.com
blogspot_outfile027	http://antifascist-calling.blogspot.com/2008_08_01_archive.html
blogspot_outfile028	http://antifascist-calling.blogspot.com/2009_05_01_archive.html
blogspot_outfile029	http://antifascist-calling.blogspot.com/2010/10/when-future-invades-our-lives-cia-funds.html
blogspot_outfile030	http://antiniworldorder.blogspot.com/2007/08/method.html
blogspot_outfile031	http://antithief.blogspot.com
blogspot_outfile032	http://apauseforbreath.blogspot.com
blogspot_outfile033	http://apexisipcamera.blogspot.com
blogspot_outfile034	http://aquaculturenamibia.blogspot.com/2015/09/fao-trainingworkshop-on-epizootic.html
blogspot_outfile035	http://arcticcompass.blogspot.com/2011_02_01_archive.html
blogspot_outfile036	http://arcticcompass.blogspot.com/2013/11/comrade-welcome-to-police-state-part-6.html
blogspot_outfile037	http://arindamcctvaccesscontrol.blogspot.com/2009/12/cctv-index.html
blogspot_outfile038	http://arindamcctvaccesscontrol.blogspot.com/2012/03/cctv-data-protection-act.html
blogspot_outfile040	http://azlyrahman-papers.blogspot.com/2005/12/41-hegemony-and-spaces-of.html
blogspot_outfile042	http://badassteachers.blogspot.com/2016/08/personalized-learning-surveillance-and.html
blogspot_outfile044	http://bauaw.blogspot.com/2016/07/bauaw-newsletter-saturday-july-9-2016.html
blogspot_outfile045	http://beatthechip.blogspot.com/2008_01_20_archive.html
blogspot_outfile046	http://behaviorist-socialist.blogspot.com/2009/03/behaviorism-is-indispensable-for.html
blogspot_outfile047	http://bestsecuritycameraofme.blogspot.com/2014_02_23_archive.html
blogspot_outfile056	http://bostonsurveillance.blogspot.com
blogspot_outfile057	http://bread-circuses-today.blogspot.com
blogspot_outfile059	http://business-security-system.blogspot.com
blogspot_outfile061	http://cahsr.blogspot.com/2009/04/homeland-security-theater.html
blogspot_outfile062	http://chewhatyoucallyourpasa.blogspot.com/2013/09/the-actual-surveillance-state.html
blogspot_outfile064	http://chycho.blogspot.com/2013/10/how-to-protect-ourselves-on-social.html
blogspot_outfile066	http://cicisocial.blogspot.com
blogspot_outfile067	http://cintronspiritualblog.blogspot.com
blogspot_outfile070	http://cisb412oct2012section2group5.blogspot.com
blogspot_outfile071	http://cisb412oct2012section2group6.blogspot.com
blogspot_outfile072	http://cisb412oct2012section2group9.blogspot.com
blogspot_outfile074	http://cloudly962.blogspot.com
blogspot_outfile075	http://coldclasscommunications.blogspot.com
blogspot_outfile078	http://conformingtoindividuality.blogspot.com
blogspot_outfile080	http://consumercal.blogspot.com/2007/10/surveillance-violates-privacy-aclu-says.html
blogspot_outfile082	http://consumercal.blogspot.com/2008_02_01_archive.html
blogspot_outfile086	http://consumercal.blogspot.com/2009/11/gps-tracking-devices-and-privacy.html
blogspot_outfile088	http://consumercal.blogspot.com/2009/12/sprint-revelations-8-million-law.html
blogspot_outfile089	http://consumercal.blogspot.com/2009_11_01_archive.html
blogspot_outfile093	http://consumercal.blogspot.com/2010/11/airport-digital-strip-searches-or.html
blogspot_outfile095	http://consumercal.blogspot.com/2010_01_01_archive.html
blogspot_outfile096	http://consumercal.blogspot.com/2010_06_01_archive.html
blogspot_outfile098	http://cornhuskeracademy.blogspot.com/2009_12_01_archive.html
blogspot_outfile100	http://country2.blogspot.com

blogspot_outfile102 <http://crimeandsurveillance.blogspot.com>
 blogspot_outfile111 <http://damanmatharoo6.blogspot.com>
 blogspot_outfile113 <http://deflem.blogspot.com/2015/01/fearofcounterterrorism.html>
 blogspot_outfile115 <http://dennismancino.blogspot.com>
 blogspot_outfile117 http://dialogic.blogspot.com/2014_01_01_archive.html
 blogspot_outfile118 <http://digigear.blogspot.com>
 blogspot_outfile119 http://dispositivodevisibilidad.blogspot.com/2007_07_01_archive.html
 blogspot_outfile120 http://downercattle.blogspot.com/2008_02_01_archive.html
 blogspot_outfile122 <http://dvr-ccd-cameras.blogspot.com>
 blogspot_outfile138 <http://elbiruniblogspotcom.blogspot.com/2015/07/ahead-of-print-real-time-microbiology.html>
 blogspot_outfile139 <http://elbiruniblogspotcom.blogspot.com/2015/11/ahead-of-print-epidemiology-of-epidemic.html>
 blogspot_outfile140 <http://electrospace.blogspot.com/2014/08/what-if-google-was-intelligence-agency.html>
 blogspot_outfile141 http://electrospace.blogspot.com/2014_03_01_archive.html
 blogspot_outfile145 <http://enjoyingthejourney.blogspot.com/2010/10/connecting-google-and-cia-dots.html>
 blogspot_outfile149 <http://erlentinadelav.blogspot.com>
 blogspot_outfile150 <http://eulawanalysis.blogspot.com/2014/12/bringing-data-protection-home-cjeu.html>
 blogspot_outfile152 <http://eunice453.blogspot.com>
 blogspot_outfile157 <http://eyespyelectronics.blogspot.com>
 blogspot_outfile159 <http://fbiorganizedstalkingchicago.blogspot.com/2015/02/what-is-organized-stalking.html>
 blogspot_outfile161 <http://fda-advocate.blogspot.com/2015/04/fda-plans-national-medical-device.html>
 blogspot_outfile169 <http://finalyearprojects--india.blogspot.com/2013/07/embedded-systems-project-titles.html>
 blogspot_outfile170 <http://fintanlawlor-dataprotection.blogspot.com>
 blogspot_outfile172 <http://fobsplace.blogspot.com>
 blogspot_outfile173 <http://frankuo.blogspot.com>
 blogspot_outfile174 <http://freecctvsecuritycamera.blogspot.com>
 blogspot_outfile176 <http://frioudshell.blogspot.com>
 blogspot_outfile179 <http://gainwealthforgetallbutself.blogspot.com>
 blogspot_outfile180 <http://gamineexpedition.blogspot.com/2009/07/data-mining-kids-and-teens-ugc-and.html>
 blogspot_outfile181 <http://gang-stalking.blogspot.com/2010/08/welcome-to-coming-surveillance-society.html>
 blogspot_outfile182 <http://getcashfrom.blogspot.com>
 blogspot_outfile183 <http://gigglesquadron.blogspot.com>
 blogspot_outfile184 <http://globale-solutions.blogspot.com/2012/11/cctv-vehicle-solution-dahua-technology.html>
 blogspot_outfile186 <http://globalrfidorder.blogspot.com>
 blogspot_outfile192 http://hissil.blogspot.com/2009/08/david-icke-guide-to-global-conspiracy_13.html
 blogspot_outfile193 <http://homesafehomes.blogspot.com>
 blogspot_outfile194 <http://hospitalsandprivacy.blogspot.com>
 blogspot_outfile195 <http://hustoknow.blogspot.com/2013/11/samsung-dvr-security-vulnerability.html>
 blogspot_outfile197 <http://idtheftsecurity.blogspot.com>
 blogspot_outfile205 <http://ifonlysingaporeans.blogspot.com/2016/03/countering-terrorism-enhancing.html>
 blogspot_outfile207 http://informedoninformation.blogspot.com/2010_10_01_archive.html
 blogspot_outfile208 <http://insan-kamil-mistik.blogspot.com>
 blogspot_outfile209 <http://internationalallawandtheinternet.blogspot.com>
 blogspot_outfile210 http://internationalsecurityjobs.blogspot.com/2012_10_01_archive.html
 blogspot_outfile211 http://internationalsecurityjobs.blogspot.com/2013_01_01_archive.html
 blogspot_outfile212 <http://internetprivacyact.blogspot.com/2012/05/internet-privacy-act-abstract.html>
 blogspot_outfile215 <http://ironwand.blogspot.com/2009/08/leo-mckinstry.html>
 blogspot_outfile217 <http://itccc.blogspot.com/2015/05/sci-diku-imageworld-digest-vol-48-issue.html>
 blogspot_outfile218 <http://itccc.blogspot.com/2016/06/sci-diku-imageworld-digest-vol-77-issue.html>
 blogspot_outfile219 <http://itccc.blogspot.com/2016/06/sci-diku-imageworld-digest-vol-78-issue.html>
 blogspot_outfile221 <http://jasontlkas.blogspot.com>
 blogspot_outfile225 <http://jobszimbabwe.blogspot.com/2016/09/employment-at-world-vision-zimbabwe.html>
 blogspot_outfile227 <http://jobszimbabwe.blogspot.com/2016/09/nutriton-assistant-zimbabwe.html>
 blogspot_outfile233 <http://just-call-me-frank.blogspot.com/2014/12/implications-of-social-networking-on.html>
 blogspot_outfile236 <http://kolkatapresents.blogspot.com/2010/09/global-law-to-kill-privacy-security.html>
 blogspot_outfile238 <http://lagmayneilryan.blogspot.com>
 blogspot_outfile240 http://lauraplanagracia.blogspot.com/2010_07_27_archive.html
 blogspot_outfile241 <http://lettheolympicgamesdieintheland.blogspot.com>
 blogspot_outfile242 <http://libertiesanddenialsof.blogspot.com>
 blogspot_outfile244 <http://lindalsmith.blogspot.com/2011/01/crime-prevention-through-environmental.html>
 blogspot_outfile245 <http://losmanifestos.blogspot.com/2005/05/el-manifiesto-del-unabomber.html>
 blogspot_outfile246 <http://loveable-eccentric.blogspot.com/2012/03/mass-media-mass-culture-and-elite.html>
 blogspot_outfile248 <http://luc-schrijvers.blogspot.com/2016/08/statewatch-news-online-30-august-2016.html>
 blogspot_outfile252 http://markbeast.blogspot.com/2009/04/prophecy-and-current-events_21.html

blogspot_outfile255 http://michelekearneyuclearwire.blogspot.com/2016_05_15_archive.html
 blogspot_outfile257 <http://moviephilippines.blogspot.com/2014/07/the-enigma-of-drones.html>
 blogspot_outfile272 <http://news33now.blogspot.com/2016/06/russian-activists-and-industry-speak.html>
 blogspot_outfile275 <http://nexusilluminati.blogspot.com/2008/01/you-will-never-be-alone.html>
 blogspot_outfile276 <http://nexusilluminati.blogspot.com/2013/12/template-for-global-prison-how.html>
 blogspot_outfile277 <http://nexusilluminati.blogspot.com/2015/07/the-matrix-deciphered-mind-control.html>
 blogspot_outfile278 <http://nexusilluminati.blogspot.com/2016/06/who-really-owns-and-controls-military.html>
 blogspot_outfile284 <http://nwoahead.blogspot.com>
 blogspot_outfile286 <http://offshoreonline.blogspot.com>
 blogspot_outfile287 <http://offshoreonline.blogspot.com/2005/07/mid-july-2005-newsletter.html>
 blogspot_outfile289 <http://onhomedsecurity.blogspot.com>
 blogspot_outfile297 http://pcwatch.blogspot.com/2015_05_01_archive.html
 blogspot_outfile298 <http://peakenergy.blogspot.com/2007/06/shockwave-rider.html>
 blogspot_outfile300 <http://permaculturepolitics.blogspot.com>
 blogspot_outfile303 <http://phspolicy.blogspot.com>
 blogspot_outfile304 <http://pippaking.blogspot.com/2013/09/really-biometrics-help-teachers-track.html>
 blogspot_outfile316 http://propheticwitnessmovementnewportbranch.blogspot.com/2015_03_01_archive.html
 blogspot_outfile317 <http://protectionplussecurity.blogspot.com>
 blogspot_outfile318 <http://rationality-rules.blogspot.com>
 blogspot_outfile321 http://realityzone-realityzone.blogspot.com/2015_12_01_archive.html
 blogspot_outfile324 <http://restoringintegritytovirginiaregistry.blogspot.com/2014/12/a-one-year-investigation-into-how-law.html>
 blogspot_outfile328 <http://rights-to-privacy.blogspot.com>
 blogspot_outfile331 <http://roshal.blogspot.com>
 blogspot_outfile332 <http://salcedorodrigo.blogspot.com>
 blogspot_outfile334 <http://santoshmilitaryscenarios.blogspot.com>
 blogspot_outfile335 <http://scintilla-spark.blogspot.com>
 blogspot_outfile338 http://securemobilephone.blogspot.com/2013_06_01_archive.html
 blogspot_outfile340 <http://securitycamerasfocus.blogspot.com>
 blogspot_outfile341 <http://securitypatrolservicesinmesquite.blogspot.com>
 blogspot_outfile342 <http://server-room-design.blogspot.com>
 blogspot_outfile347 <http://sitihajarismail-tinta.blogspot.com>
 blogspot_outfile348 <http://sitsshow.blogspot.com/2015/07/the-matrix-deciphered-mind-control.html>
 blogspot_outfile349 <http://sitsshow.blogspot.com/2016/05/Who-Really-Owns-and-Controls-the-Military-Industrial-Complex-and-What-Are-They-Doing-An-Extensive-Research-Report.html>
 blogspot_outfile358 <http://standalone-dvrs.blogspot.com>
 blogspot_outfile359 <http://stanvanhoucke.blogspot.com/2015/03/paul-heeft-een-nieuwe-reactie-op-je.html>
 blogspot_outfile361 http://strategicstudyindia.blogspot.com/2016_03_28_archive.html
 blogspot_outfile363 <http://surveillancespot.blogspot.com>
 blogspot_outfile364 <http://surveillantidentity.blogspot.com>
 blogspot_outfile365 http://surveillantidentity.blogspot.com/2011_04_01_archive.html
 blogspot_outfile368 <http://szaboservices.blogspot.com/2011/07/anonymous-shares-90000-military-email.html>
 blogspot_outfile373 <http://tcnwo.blogspot.com>
 blogspot_outfile374 <http://tcnwo.blogspot.com/p/national-id-cards.html>
 blogspot_outfile377 <http://tekicctvreport.blogspot.com>
 blogspot_outfile378 <http://terry-young.blogspot.com>
 blogspot_outfile379 <http://the-malaysian.blogspot.com/2008/01/why-malaysia-is-ranked-worst-country-in.html>
 blogspot_outfile381 <http://thecomunitariantrap.blogspot.com>
 blogspot_outfile382 <http://thecyberwilderness.blogspot.com>
 blogspot_outfile383 <http://theflamingmoderate.blogspot.com>
 blogspot_outfile384 <http://theloadedxraygun.blogspot.com>
 blogspot_outfile385 <http://thenwwrdor.blogspot.com/2013/04/14-incredibly-creepy-surveillance.html>
 blogspot_outfile391 http://thewesterner.blogspot.com/2012_06_17_archive.html
 blogspot_outfile393 http://tomsfbidenity.blogspot.com/2013_10_01_archive.html
 blogspot_outfile394 http://tonymusings.blogspot.com/2014_02_01_archive.html
 blogspot_outfile401 <http://univercia22.blogspot.com/2010/05/combined-us-israeli-espionage-of-global.html>
 blogspot_outfile402 <http://uonica.blogspot.com>
 blogspot_outfile404 <http://viclivingston.blogspot.com/2011/12/u.html>
 blogspot_outfile405 <http://viclivingston.blogspot.com/2012/01/thugocracy-u.html>
 blogspot_outfile406 <http://viclivingston.blogspot.com/2012/01/top-brass-piss-on-marines-as-they-cover.html>
 blogspot_outfile408 <http://wanabidiiblog.blogspot.com/2012/10/wanabidii-vision-2030-is-serious.html>
 blogspot_outfile409 <http://wantotellyou.blogspot.com>
 blogspot_outfile410 http://warningsforamericasfuture.blogspot.com/2015_05_01_archive.html
 blogspot_outfile411 <http://waterloowatchmen.blogspot.com>

blogspot_outfile412 <http://whysurveil.blogspot.com>
 blogspot_outfile414 <http://wirelesssecuritycamerasforhome.blogspot.com>
 blogspot_outfile415 <http://womenincrimeink.blogspot.com/2011/02/whats-in-face-biometrics-and-facial.html>
 blogspot_outfile416 <http://word-up-studies.blogspot.com/2009/12/shadow-government-by-grant-jeffrey.html>
 blogspot_outfile417 <http://worldunderwatch.blogspot.com/2010/08/chinas-government-no-stranger-to-big.html>
 blogspot_outfile418 <http://worldunderwatch.blogspot.com/2010/08/new-delhi-police-look-to-facebook-to.html>
 blogspot_outfile419 <http://worldunderwatch.blogspot.com/2010/09/us-federal-law-enforcement-agencies.html>
 blogspot_outfile420 <http://worldunderwatch.blogspot.com/p/camera-laws.html>
 blogspot_outfile421 <http://worldunderwatch.blogspot.com/p/site-mission.html>
 blogspot_outfile422 <http://zaf512blog.blogspot.com>
 blogspot_outfile424 <http://zoompad.blogspot.com/2013/09/the-prevention-of-terrorism-bill.html>
 blogspot_outfile426 https://bikersofamerica.blogspot.com/2014_07_01_archive.html
 blogspot_outfile428 https://bikersofamerica.blogspot.com/2016_03_01_archive.html
 blogspot_outfile430 <https://mkaguzi.blogspot.com/2016/05/malaria-surveillance-monitoring.html>
 blogspot_outfile433 https://spybusters.blogspot.com/2013_06_01_archive.html
 blogspot_outfile434 <https://spybusters.blogspot.com/2014/11/73012-unsecured-security-cameras-you.html>
 blogspot_outfile435 https://spybusters.blogspot.com/2015_12_01_archive.html
 blogspot_outfile436 <https://stevenwarran.blogspot.com/2012/07/it-wasnt-muslims.html>
 blogspot_outfile437 <https://stevenwarran.blogspot.com/2013/02/paul-mozer-salomon-bond-trader.html>
 typepad_outfile003 <http://amberhawk.typepad.com/amberhawk/2010/03/do-local-authorities-spend-nearly-half-a-billion-pounds-on-ineffective-cctv.html>
 typepad_outfile004 <http://amberhawk.typepad.com/amberhawk/2010/08/data-protection-and-surveillance-swopping-the-speed-camera-for-anpr.html>
 typepad_outfile006 <http://amberhawk.typepad.com/amberhawk/2010/08/swimming-in-the-surveillance-pool-the-real-privacy-problems-with-ripa.html>
 typepad_outfile007 <http://amberhawk.typepad.com/amberhawk/2010/10/index.html>
 typepad_outfile008 <http://amberhawk.typepad.com/amberhawk/2010/10/project-champion-report-misses-the-target-when-discussing-cctv-surveillance-privacy-and-data-protection.html>
 typepad_outfile010 <http://amberhawk.typepad.com/amberhawk/2012/10/mobile-cctv-cars-used-by-councils-can-breach-data-protection-law-and-ripa.html>
 typepad_outfile012 <http://amberhawk.typepad.com/amberhawk/2013/08/index.html>
 typepad_outfile014 <http://amberhawk.typepad.com/amberhawk/data-protection/page/13>
 typepad_outfile015 <http://amberhawk.typepad.com/amberhawk/data-protection/page/14>
 typepad_outfile016 <http://amberhawk.typepad.com/amberhawk/data-protection/page/19>
 typepad_outfile017 <http://amberhawk.typepad.com/amberhawk/data-protection/page/2>
 typepad_outfile018 <http://amberhawk.typepad.com/amberhawk/data-protection/page/20>
 typepad_outfile020 <http://amberhawk.typepad.com/amberhawk/data-protection/page/22>
 typepad_outfile023 <http://amberhawk.typepad.com/amberhawk/data-protection/page/30>
 typepad_outfile027 <http://amberhawk.typepad.com/amberhawk/news/page/13>
 typepad_outfile031 <http://amberhawk.typepad.com/amberhawk/other-information-law/page/10>
 typepad_outfile033 <http://amberhawk.typepad.com/amberhawk/other-information-law/page/2>
 typepad_outfile038 <http://amberhawk.typepad.com/amberhawk/page/21>
 typepad_outfile044 <http://calsaga.typepad.com/calsaga/2007/08/daily-private-2.html>
 typepad_outfile049 http://chutzpah.typepad.com/slow_movement/analytics
 typepad_outfile058 http://elainemeinelsupkis.typepad.com/war_and_peace/page/2
 typepad_outfile059 <http://fraser.typepad.com/socialtech/communities>
 typepad_outfile094 http://sentencing.typepad.com/sentencing_law_and_policy/2012/12/index.html
 typepad_outfile101 http://sentencing.typepad.com/sentencing_law_and_policy/technocorrections
 typepad_outfile126 <http://tommytoy.typepad.com/tommy-toy-pbt-consultin/consumer-rights>
 typepad_outfile131 <http://tommytoy.typepad.com/tommy-toy-pbt-consultin/security-systems-and-devices>
 typepad_outfile137 <http://whirledview.typepad.com/whirledview/science/index.html>
 wordpress_outfile001 <https://aayaminfotech.wordpress.com>
 wordpress_outfile002 <https://actnowtraining.wordpress.com/2015/03/03/cctv-surveillance-getting-it-right>
 wordpress_outfile003 <https://actnowtraining.wordpress.com/2015/10/01/cctv-and-the-law>
 wordpress_outfile004 <https://aithyia.wordpress.com>
 wordpress_outfile006 <https://alethonews.wordpress.com/2014/11/22/security-in-the-wests-client-states>
 wordpress_outfile007 <https://alethonews.wordpress.com/category/subjugation-torture>
 wordpress_outfile008 <https://alethonews.wordpress.com/tag/human-rights>
 wordpress_outfile009 <https://alethonews.wordpress.com/tag/nsa>
 wordpress_outfile013 <https://ashjamie.wordpress.com/2016/09/06/surveillance-liked-your-facebook-status>
 wordpress_outfile014 <https://asyafaat.wordpress.com/2013/06/26/prism-program-pengintaian-internet-as>
 wordpress_outfile015 <https://ausltechlaw.wordpress.com/2012/10/01/obmaces-rhazzlet-cctv-adoption-in-public-places-vs-privacy-matters>

wordpress_outfile016 <https://aveasia.wordpress.com/2015/01/23/genetec-to-support-new-microsoft-cloud-for-government>

wordpress_outfile017 <https://axiomamuse.wordpress.com/category/international-association-of-chiefs-of-police>

wordpress_outfile019 <https://axiomamuse.wordpress.com/tag/intelligence-led-policing>

wordpress_outfile023 <https://berarespectingchildren.wordpress.com/2014/08/21/the-militarisation-of-education-troops-to-teachers-and-the-implications-for-initial-teacher-education-and-race-equality>

wordpress_outfile026 <https://blackclawspy.wordpress.com/2010/02/11/spies-in-the-ointment-the-israeli-espionage-of-global-communications>

wordpress_outfile027 <https://brandkemistry.wordpress.com/category/thursday-job-posts>

wordpress_outfile028 <https://brettideas.wordpress.com/2014/04/03/privacy-and-disclosure-an-online-balancing-act>

wordpress_outfile038 <https://etisalatyellowpages.wordpress.com>

wordpress_outfile040 <https://findmyfriendsapplication.wordpress.com>

wordpress_outfile043 <https://gachekegachihi.wordpress.com>

wordpress_outfile044 <https://gfbertini.wordpress.com/2013/09/29/spying-on-democracy-government-surveillance-corporate-power-and-public-resistance>

wordpress_outfile057 <https://haddensecurity.wordpress.com/category/special-dossiers-ingles>

wordpress_outfile058 <https://hellotheinternet.wordpress.com/2016/10/06/the-news-for-yous>

wordpress_outfile059 <https://homelessholocaust.wordpress.com/2011/12/12/orwellian-laws-for-poor>

wordpress_outfile062 <https://hrdepartmentireland.wordpress.com/category/cctv>

wordpress_outfile064 <https://iassupermitra.wordpress.com>

wordpress_outfile069 <https://introlex.wordpress.com>

wordpress_outfile071 <https://jagadees.wordpress.com/2016/07/26/fukushima-daiichi-nuclear-accident-is-wake-up-call-for-u-s>

wordpress_outfile072 <https://jelkingtonblog.wordpress.com>

wordpress_outfile074 <https://johnib.wordpress.com/2016/09/17/germany-merkel-faces-further-losses-in-regional-elections>

wordpress_outfile075 <https://johnnicholson.wordpress.com>

wordpress_outfile079 <https://kevinclementspeaceandconflict.wordpress.com>

wordpress_outfile080 <https://kittysjones.wordpress.com/2013/09/18/we-must-oppose-the-coalitions-mandatory-national-ids-and-biometric-systems>

wordpress_outfile082 <https://knownuclearwaste.wordpress.com/2016/05/24/fukushima-nuclear-accident-is-wake-up-call-for-us-to-improve-monitoring-of-spent-fuel-pools>

wordpress_outfile083 <https://konnichiwadotcom.wordpress.com/privacy-issues-in-the-digital-age>

wordpress_outfile084 <https://lastfreelance.wordpress.com/2015/08/07/comparative-analysis>

wordpress_outfile086 <https://lazokoroji.wordpress.com/2013/01/04/a-paper-on-the-international-ship-and-port-facility-security-code-isps>

wordpress_outfile087 <https://libertyhereandnow.wordpress.com>

wordpress_outfile089 <https://liturgical.wordpress.com/2008/03/11/cdcs-report-on-teen-girls-with-stds-considered-with-no-2-song-on-itunes>

wordpress_outfile091 <https://lyrebirdpolitik.wordpress.com>

wordpress_outfile096 <https://medtextfree.wordpress.com/2011/04/06/5-4-applications-of-information-systems-to-public-health>

wordpress_outfile097 <https://medtextfree.wordpress.com/2011/05/24/6-15-microcomputers-the-internet-and-epidemiology>

wordpress_outfile098 <https://medtextfree.wordpress.com/page/73>

wordpress_outfile099 <https://medtextfree.wordpress.com/page/75>

wordpress_outfile100 <https://mintehbaa.wordpress.com/2011/12/22/cctv-is-a-major-component-of-many-physical-security-systems-recent-technological-developments-have-included-colour-solid-state-ccd-cameras-and-image-intensification-discuss-the-factors-to-be-considered>

wordpress_outfile106 <https://moorbey.wordpress.com/2015/09/11/coming-home-to-roost-american-militarism-war-kulture-and-police-brutality>

wordpress_outfile109 <https://neoconopticon.wordpress.com/tag/finmeccanica>

wordpress_outfile110 <https://neoconopticon.wordpress.com/tag/frontex>

wordpress_outfile111 <https://news247worldpressuk.wordpress.com/2016/08/26/country-reports-on-terrorism-2015>

wordpress_outfile112 <https://niqnaq.wordpress.com/2013/08/31/spoof-nsa-site>

wordpress_outfile113 <https://noelthecyberlawyer.wordpress.com>

wordpress_outfile118 <https://nwoobserver.wordpress.com/tag/big-brother/page/3>

wordpress_outfile119 <https://oceanflynn.wordpress.com/2007/11/30/media-objectivity-in-canada-a-timeline-of-social-events>

wordpress_outfile122 <https://oksafe.wordpress.com/category/surveillance-society>

wordpress_outfile123 <https://oksafe.wordpress.com/tag/surveillance>

wordpress_outfile125 <https://orderoftruth.wordpress.com/2013/06/24/why-law-abiding-citizens-should-be-worried-about-the-developing-surveillance-state-video>

wordpress_outfile126 <https://paulbernal.wordpress.com/2012/08>

wordpress_outfile127 <https://paulbernal.wordpress.com/2013/03/25/the-snoopers-charter-we-need-a-new-consultation>

wordpress_outfile129 <https://paulbernal.wordpress.com/2014/02/06/communications-surveillance-a-miscast-debate>

wordpress_outfile130 <https://paulbernal.wordpress.com/2014/04/08/data-retention-fighting-for-privacy>

wordpress_outfile136 <https://paulbernal.wordpress.com/category/internet/page/4>

wordpress_outfile137 <https://paulbernal.wordpress.com/category/privacy-international/page/2>

wordpress_outfile138 <https://paulbernal.wordpress.com/category/privacy/page/3>

wordpress_outfile139 <https://paulbernal.wordpress.com/category/privacy/page/6>

wordpress_outfile142 <https://paulbernal.wordpress.com/category/tracking>

wordpress_outfile144 <https://paulbernal.wordpress.com/tag/data-protection-2>

wordpress_outfile152 <https://peepbeep.wordpress.com/category/surveillance>

wordpress_outfile155 <https://peterbdunn.wordpress.com/2014/01/26/echelon-and-high-tech-intelligence-records-for-you>

wordpress_outfile159 <https://piava.wordpress.com/category/surveillance>

wordpress_outfile160 <https://pinkindustry.wordpress.com/the-quilliam-foundation-the-media>

wordpress_outfile163 <https://policedeviance.wordpress.com/2015/02/12/body-cameras-a-new-extension-of-police-visibility>

wordpress_outfile164 <https://policedeviance.wordpress.com/2015/02/page/5>

wordpress_outfile166 <https://policedeviance.wordpress.com/tag/surveillance>

wordpress_outfile167 <https://pressall.wordpress.com/tag/twitter/page/2>

wordpress_outfile169 <https://privacynewshighlights.wordpress.com/2011/06/29/16-31-may-2011>

wordpress_outfile170 <https://privacynewshighlights.wordpress.com/2014/06/12/15-28-february-2014>

wordpress_outfile171 <https://privacynewshighlights.wordpress.com/2014/12>

wordpress_outfile172 <https://privacynewshighlights.wordpress.com/2015/07/07/16-30-june-2015>

wordpress_outfile173 <https://privacynewshighlights.wordpress.com/2015/11/27/01-15-november-2015>

wordpress_outfile174 <https://privacynewshighlights.wordpress.com/2016/06/20/10-17-june-2016>

wordpress_outfile175 <https://privacynewshighlights.wordpress.com/2016/07/06/25-june-04-july-2016>

wordpress_outfile180 <https://raymondpronk.wordpress.com/2014/01/04/national-security-agency-nsa-trailblazer-the-program-stellar-wind-prism-xkeyscore-photos-slides-videos>

wordpress_outfile183 <https://realagenda.wordpress.com/tag/national-security-agency>

wordpress_outfile186 <https://researchedsolution.wordpress.com/2013/09/page/2>

wordpress_outfile188 <https://researchhumantrafficking.wordpress.com/2014/09/03/the-trafficking-technology-nexus>

wordpress_outfile192 <https://rortybomb.wordpress.com/2011/05/26/balkin-on-the-constitution-in-the-national-surveillance-state>

wordpress_outfile198 <https://simplysociology.wordpress.com/tag/surveillance-society>

wordpress_outfile200 <https://smdawes.wordpress.com/2014/04/22/article-press-freedom-privacy-and-the-public-sphere>

wordpress_outfile209 <https://sohereandnow.wordpress.com/2013/02/02/activist-post-by-nicholas-post>

wordpress_outfile210 <https://sophielmcadam.wordpress.com/2010/11/13/you-have-the-right-to-remain-silent>

wordpress_outfile211 <https://spectregroup.wordpress.com/2008/04/01/r-u-in-jail>

wordpress_outfile212 <https://spfaust.wordpress.com/category/privacy>

wordpress_outfile213 <https://spideytrsr.wordpress.com>

wordpress_outfile216 <https://thafcc.wordpress.com/2015/04/03/the-fbi-has-its-own-secret-brand-of-malware>

wordpress_outfile217 <https://thdblog.wordpress.com/category/public-private-partnerships>

wordpress_outfile219 <https://thecommune.wordpress.com/2010/09/19/cuba-the-island-of-happiness>

wordpress_outfile224 <https://theoryculturesociety2015.wordpress.com>

wordpress_outfile227 <https://thinkingshift.wordpress.com/2007/12/page/2>

wordpress_outfile228 <https://thinkingshift.wordpress.com/2008/05/09/how-to-spy-on-british-motorists>

wordpress_outfile229 <https://thinkingshift.wordpress.com/category/cctv>

wordpress_outfile231 <https://thinkingshift.wordpress.com/tag/united-kingdom>

wordpress_outfile237 <https://totalsecuritywarehouse.wordpress.com>

wordpress_outfile239 <https://triumphedbetime.wordpress.com/2014/12/22/funlux-8ch-960h-qr-code-easy-setup-video-security>

wordpress_outfile241 <https://ubisurv.wordpress.com/2009/02/03/more-cctv-stories-from-jerusalem-to-cambridge>

wordpress_outfile242 <https://ubisurv.wordpress.com/2009/03/02/how-many-cameras-are-there-in-britain>

wordpress_outfile243 <https://ubisurv.wordpress.com/2009/03/04/a-sane-response-on-the-number-of-cameras>

wordpress_outfile244 <https://ubisurv.wordpress.com/2009/08/10/china-the-ultimate-surveillance-state>

wordpress_outfile245 <https://ubisurv.wordpress.com/2009/10/16/surveillance-image-of-the-week-2-camera-catches-man-stealing-camera>

wordpress_outfile246 <https://undercoverinfo.wordpress.com/2015/07/02/a-journalists-whistleblowers-brief-guide-to-sabotaging-gagging-laws>

wordpress_outfile247 <https://wakinguporwell.wordpress.com/2011/08/16/basic-psychology-of-surveillance>

wordpress_outfile248 <https://warriorpublications.wordpress.com/2013/12/31/counter-insurgency-stratfor-monitors-and-studies-social-movements>

wordpress_outfile249	https://warriorpublications.wordpress.com/2014/04/25/fbi-building-biometric-database-containing-face-scans-of-millions-of-americans
wordpress_outfile251	https://waveupdatessite.wordpress.com/2016/09/09/liberia-senior-anti-corruption-advisor-enforcement-legal-professional-development-and-anti-corruption-program-lpac-liberia
wordpress_outfile252	https://wdsi.wordpress.com/2012/05/26/terrorism-bookshelf-top-150-books-on-terrorism-and-counterterrorism
wordpress_outfile253	https://welections.wordpress.com/page/2
wordpress_outfile255	https://whereismydata.wordpress.com/category/3-data/cctv/page/3
wordpress_outfile260	https://zainindo.wordpress.com/2009/03/25/cctv-wiki